



FACULTAD DE CIENCIAS
Departamento de Biología Molecular

**BASES MOLECULARES DE LA REPLICACIÓN
DEL CORONAVIRUS TGEV**

CARMEN GALÁN AVELLA



CENTRO NACIONAL DE BIOTECNOLOGÍA

Madrid, 2007



FACULTAD DE CIENCIAS
Departamento de Biología Molecular

BASES MOLECULARES DE LA REPLICACIÓN DEL CORONAVIRUS TGEV

CARMEN GALÁN AVELLA

Madrid, Enero de 2007

UNIVERSIDAD AUTÓNOMA DE MADRID

FACULTAD DE CIENCIAS

Departamento de Biología Molecular

**BASES MOLECULARES DE LA REPLICACIÓN
DEL CORONAVIRUS TGEV**

Memoria presentada por Carmen Galán Avella
para optar al grado de Doctor en Ciencias
por la Universidad Autónoma de Madrid

Madrid, Enero de 2007



El trabajo que se describe en esta memoria ha sido realizado en el Centro Nacional de Biotecnología (CNB, CSIC) bajo la codirección de los Drs. Fernando Almazán y Luis Enjuanes.

Madrid, Enero de 2007

Fdo. Fernando Almazán

Fdo. Luis Enjuanes

AGRADECIMIENTOS

Gracias a todos los que me habéis ayudado a que hoy esté escribiendo la última página de la tesis:

A los directores de este trabajo, el profesor de investigación Luis Enjuanes y el doctor Fernando Almazán, porque he recibido de vosotros todo el apoyo profesional y un excelente trato personal. Gracias Luis, por tu confianza y por todas las facilidades y medios con que has contribuido a mi formación dentro y fuera de tu laboratorio. Fernando, gracias por tu tiempo e implicación, por la tolerancia de todos los días.

Agradezco la financiación obtenida del Ministerio de Ciencia y Tecnología mediante la concesión de una beca FPI asociada al proyecto BIO2001-1699.

Mi mejor descubrimiento, las personas que fueron y son el 114: a los doctores Jose Manuel S., David E., Sara A., y mi querida Cristina R., os deseo lo mejor. Diana, me ha “prestado” mucho coincidir contigo. Gracias Verónica, por tu cariño.

A la Dra. Isabel Sola por tu experiencia siempre disponible, admiro tu integridad científica y humana. A M^a Carmen Capiscol por tu franqueza y por tu lado mejor siempre. A la Dra. Sonia Zúñiga por tu cercanía y tu valiosa ayuda informática y científica.

Aitor, eres un compañero excelente (perdóname el secuestro de tu ordenador). A Jose Luis M., por los “mínimos” detalles que hacen grandes a las personas. Gracias a la amabilidad de Marga y al buen humor de los Drs. Kike y Juan. Al Dr. Javier Ortego, por cultivar las relaciones personales dentro y fuera del laboratorio.

Al Dr. Carlos S., a Marta L., Jazmina y Beatriz... a todos y cada uno, gracias por vuestro apoyo a pie de trinchera.

A los compañeros del Departamento de Biología Molecular que repartís ánimos y buenos-días por los pasillos. Marta G., Inma, Miguel, Rubén... os echo en falta.

Al servicio de proteómica del CNB, en especial a Silvia Juárez, por su excelente trabajo, consulta técnica y trato personal, así da gusto recortar proteínas.

Gracias al Dr. Sasha Akoulitchev de la Universidad de Oxford por su acogida y ayuda durante mi estancia en su laboratorio. Al Dr. Francisco Iborra y familia por su hospitalidad. Al Dr. Benjamin Thomas por poner a mi disposición su experiencia en espectrometría de masas y a todo el laboratorio de Bioquímica y Dinámica Nuclear, porque hicisteis más cálido mi invierno en Inglaterra.

Un recuerdo para mis profesores del colegio Auseva de Oviedo, Carmen Fuertes y Jose F. Castro, cuyo entusiasmo me enganchó a la ciencia.

Otro para los buenos amigos que me dejó Bioquímica, Marta M., Antonio, Óscar, Elise...

Gracias al profesor de investigación Antonio Villalobo por iniciarme en la investigación y animarme a continuar.

A los Amigos que me habéis vivido más de cerca estos años, gracias por haber sido tantas veces mi toma de tierra y mi salida de emergencia (ya fuera a la Patagonia o a Cobeña). A Tomás por tu amistad incondicional, a María, alegría a domicilio en cualquier momento, a Pilar por tenerte tan cerca, estés donde estés. A Laura y Jose, por estar ahí en horas altas y bajas, y a Jesús por la mejor conversación en cercanía(s).

A todos los que me habéis hecho querer Alcalá y necesitaríais capítulo a parte.

Gracias a Eduardo Suárez, por escribir una página especial.

A Isabel González, por ser como eres, y a Paco Barrios, que tienes un sitio reservado en mi corazón. Gracias a los dos por seguir estando.

Gracias a mi familia por tanto cariño, a mi tía Loly, que encuentras el tiempo que no tienes para que no me falten tus detalles, a mis hermanos Jesús y Luis, lo mejor que tengo... y a la memoria de mi abuela Carmen, ella sabe por qué.

Por último, y en primer lugar, a mis padres Luis y Maria del Carmen. Os debo todo, sois mi tierra firme. A vosotros está dedicada cada palabra de esta tesis.

*El hombre encuentra a Dios detrás de
cada puerta que la ciencia logra abrir*

Albert Einstein

A mis padres

ABREVIATURAS

aa	Aminoácido
APN	Aminopeptidasa N, CD13
ATCC	<i>American type culture collection</i> (colección americana de cultivos tipo)
BAC	<i>Bacterial artificial chromosome</i> (cromosoma artificial bacteriano)
BGH	<i>Bovine growth hormone</i> (hormona de crecimiento bovina)
BW	<i>Binding/washing buffer</i> (tampón de unión y lavado)
CapLC	<i>Capillary liquid chromatography</i> (cromatografía líquida capilar)
CAT	Cloranfenicol acetil transferasa
cDNA	Ácido desoxirribonucleico complementario
3CL	<i>Chemotrypsin like-proteinase</i> (proteínasa similar a la quimiotripsina o proteinasa principal, M ^{pro} , similar a la proteinasa 3C de picornavirus)
CMV	Citomegalovirus
CS	<i>Consensus sequence</i> (secuencia consenso para la transcripción)
CoV	Coronavirus
DEPC	Diethylpirocarbonato
DIG	Digoxigenina
DMEM	<i>Dulbecco's modified eagle medium</i> (medio <i>Eagle</i> modificado por Dulbecco)
DMSO	Dimetil sulfóxido
DNA	<i>Deoxyribonucleic acid</i> (ácido desoxirribonucleico, ADN)
DO	Densidad óptica
DTT	1,4-Ditiotreitol
EAV	<i>Equine arteritis virus</i> (virus de la arteritis equina)
EDTA	Ácido etilén diaminotetraacético
EGFP	<i>Enhanced green fluorescence protein</i> (proteína verde fluorescente intensificada)
ESI	<i>Electrospray ionization source</i> (fuente de ionización por electrospray)
EtOH	Etanol
GFP	<i>Green fluorescence protein</i> (proteína verde fluorescente)
GUS	β-Glucuronidasa
HDV	<i>Hepatitis delta virus</i> (virus de la hepatitis delta)
hdi	Horas después de la infección
hdt	Horas después de la transfección
HeLa	Línea celular de carcinoma cervical humano
HEPES	Ácido N-2-hidroxietilpiperacina-N'-2-etanosulfónico
HnRNP	<i>Heterogeneous nuclear ribonucleoprotein</i> (ribonucleoproteína heterogénea nuclear)
HPLC	<i>High performance liquid chromatography</i> (cromatografía líquida de alta resolución)
HSP	<i>Heat shock protein</i> (proteína inducida por choque térmico)
IBV	<i>Avian infectious bronchitis virus</i> (virus de la bronquitis infecciosa aviar)
IRES	<i>Internal ribosome entry site</i> (sitio de entrada interna del ribosoma)
kb	Kilobase
kDa	Kilodalton
LB	Luria-Bertani
MALDI	<i>Matrix-assisted laser desorption/ionization</i> (desorción/ionización por láser asistida por matriz)
MCS	<i>Multicloning site</i> (sitio múltiple de clonaje)
Mdi	Multiplicidad de infección
MHV	<i>Mouse hepatitis virus</i> (virus de la hepatitis murina)
mRNA	<i>Messenger RNA</i> (ácido ribonucleico mensajero)
MS	<i>Mass spectrometry</i> (espectrometría de masas)
MS/MS	<i>Tandem mass spectrometry</i> (espectrometría de masas en rondas sucesivas)
NFAR	<i>Nuclear factor associated with dsRNA</i> (proteínas que unen RNA de cadena doble)
nsp	<i>Nonstructural protein</i> (proteína no estructural)
nt	Nucleótido
ORF	<i>Open reading frame</i> (fase abierta de lectura)
PABP	<i>Polyadenine binding protein</i> (proteína de unión a tramos de poliadenina)
PAGE	<i>Polyacrylamide gel electrophoresis</i> (electroforesis en gel de poliacrilamida)

Abreviaturas

pAPN	<i>Porcine aminopeptidase N</i> (aminopeptidasa N porcina)
pAPN-BHK	<i>pAPN-baby hamster kidney cells</i> (línea celular de riñón de hámster que expresa la aminopeptidasa N porcina)
PBS	Tampón fosfato salino
PCR	<i>Polymerase chain reaction</i> (reacción en cadena de la polimerasa)
PLP	<i>Papain like proteinase</i> (proteínasa similar a la papaína)
poliA	Tramo de poliadeninas, An
poliU	Tramo de poliuridinas, Un
PTB	<i>Polypirimidine-tract binding protein</i> (proteína de unión a tramos de polipirimidina)
PUR46-MAD	Virus TGEV, clon Purdue 46-Madrid
PUR46-C11	Virus TGEV, clon Purdue 46-C11
Q-RT-PCR	<i>Quantitative-reverse transcription-polymerase chain reaction</i> (transcripción reversa seguida de reacción en cadena de la polimerasa cuantitativa)
QUAD	Cuadripolo, un tipo de analizador de masas
RNA	<i>Ribonucleic acid</i> (ácido ribonucleico, ARN)
rRNA	<i>Ribosomal ribonucleic acid</i> (ácido ribonucleico ribosomal)
RS	<i>Reverse sense</i> (sentido reverso)
RT	<i>Reverse transcription/transcriptase</i> (transcripción inversa; también transcriptasa reversa)
rTGEV	<i>Recombinant transmissible gastroenteritis virus</i> (virus TGEV recombinante)
SDS	Dodecilsulfato sódico
SARS	<i>Severe acute respiratory syndrome</i> (síndrome respiratorio agudo y grave)
ssDNA	<i>single stranded DNA</i> (DNA de cadena sencilla)
ST	<i>Swine testis</i> (línea celular de testículo de cerdo)
STF	Suero de ternera fetal
TAE	Tampón Tris-acetato-EDTA
TBE	Tampón Tris-borato-EDTA
TE	Tampón Tris-EDTA
TGEV	<i>Transmissible gastroenteritis virus</i> (virus de la gastroenteritis porcina transmisible)
Tm	Temperatura de fusión
TOF	<i>Time of flight</i> (tiempo de vuelo, un tipo de analizador de masas)
tRNA	<i>Transfer ribonucleic acid</i> (ácido ribonucleico de transferencia)
TRS	<i>Transcription-regulating sequence</i> (secuencia reguladora de la transcripción)
UFP	Unidades formadoras de placa
UTR	<i>Untranslated region</i> (región no traducida)
cUTR	<i>Complementary UTR</i> (región no traducida, secuencia complementaria)
UV	Ultravioleta
VS	<i>Virus sense</i> (sentido viral)

I. ÍNDICE

I. ÍNDICE

II. RESUMEN EN INGLÉS.....	1
III. INTRODUCCIÓN.....	3
1. REPLICACIÓN DE VIRUS RNA DE CADENA SENCILLA Y POSITIVA.....	3
2. CARACTERÍSTICAS GENERALES DE LOS CORONAVIRUS.....	3
3. EL VIRUS DE LA GASTROENTERITIS PORCINA TRANSMISIBLE.....	5
3.1. Estructura y composición del virión.....	6
3.2. Estructura del genoma y expresión génica.....	8
3.3. Ciclo infectivo viral.....	10
4. GENÉTICA REVERSA EN CORONAVIRUS.....	11
4.1. Sistemas dependientes de virus complementador.....	12
4.2. Generación de clones infectivos de coronavirus.....	13
4.3. Los coronavirus como vectores de expresión.....	14
5. SÍNTESIS DE RNA EN CORONAVIRUS.....	15
5.1. Transcripción	16
5.2. Replicación	18
5.2.1. Señales en <i>cis</i> necesarias para la replicación	18
5.2.2. Estructuras de RNA implicadas en la replicación.....	20
6. LA REPLICASA DE CORONAVIRUS	20
7. PROTEÍNAS CELULARES IMPLICADAS EN LA REPLICACIÓN DE CORONAVIRUS	23
IV. OBJETIVOS.....	27
V. MATERIALES Y MÉTODOS.....	29
1. CÉLULAS EUCARIÓTICAS.....	29
1.1. Líneas celulares.....	29
1.2. Cultivo de células eucarióticas.....	29
2. VIRUS.....	29
2.1. Aislados virales.....	29
2.2. Crecimiento y titulación del virus TGEV.....	30

3. BACTERIAS.....	30
3.1. Cepas y cultivos de bacterias.....	30
3.2. Preparación de bacterias competentes.....	30
3.3. Transformación de bacterias.....	31
4. MANIPULACIÓN Y ANÁLISIS DEL DNA.....	31
4.1. Plásmidos de clonaje.....	31
4.2. Manipulación y purificación de plásmidos.....	32
4.3. Enzimas de restricción y modificación de DNA.....	32
4.4. Amplificación del DNA mediante PCR.....	32
4.5. Electroforesis de DNA en geles de agarosa.....	33
5. CONSTRUCCIÓN Y MUTAGÉNESIS DE PLÁSMIDOS.....	33
5.1. Mutagénesis de cDNAs del genoma del TGEV.....	33
5.2. Modificación y mutagénesis de minigenomas derivados del TGEV.....	34
5.3. Clonaje del dominio proteinasa PLP1 y del sustrato N-terminal de la replicasa	36
5.4. Generación de un plásmido no replicativo derivado del clon infectivo del TGEV para optimizar la eficiencia de transfección de DNAs de gran tamaño....	37
5.5. Construcción de replicones derivados del genoma del TGEV.....	37
5.6. Mutagénesis del gen de la proteína N en el replicón del virus Sindbis	40
5.7. Construcción de plásmidos para la transcripción <i>in vitro</i> de los extremos 5' y 3' del genoma del TGEV.....	41
6. MANIPULACIÓN Y ANÁLISIS DEL RNA.....	43
6.1. Extracción y purificación del RNA intracelular.....	43
6.2. Análisis del RNA mediante RT-PCR.....	43
6.3. Cuantificación del RNA mediante RT-PCR a tiempo real.....	44
7. TRANSFECCIÓN DE DNA.....	45
7.1. Métodos utilizados para optimizar la transfección de BACs en células de mamífero.....	45
7.2. Evaluación de la eficiencia de transfección mediante citometría de flujo.....	46
8. MICROSCOPIA DE FLUORESCENCIA.....	46
9. METODOLOGÍA PARA LOS ESTUDIOS DE AMPLIFICACIÓN DE MINIGENOMAS POR VIRUS COMPLEMENTADOR.....	47
9.1. Transcripción <i>in vitro</i> de minigenomas derivados del TGEV.....	47
9.2. Transfección de células infectadas con minigenomas.....	48
9.3. Análisis del rescate de minigenomas a lo largo de los pases.....	48

10. TRANSFECCIÓN Y RESCATE DE VIRUS A PARTIR DE CLONES INFECTIVOS.....	48
11. ENSAYO DE PROCESAMIENTO EN <i>TRANS</i> DE LA REPLICASA POR LA PROTEASA VIRAL PLP1.....	49
12. CARACTERIZACIÓN DE REPLICONES DERIVADOS DEL GENOMA DEL TGEV.....	50
12.1. Análisis de la actividad de los replicones mediante RT-PCR y Q-RT-PCR.....	50
12.2. Determinación de la actividad β -glucuronidasa.....	50
12.3. Análisis del efecto citopático inducido por los replicones diferentes líneas celulares	51
12.4. Generación de células BHK-pAPN transfectadas con el replicón del virus Sindbis expresando diferentes mutantes del gen N.....	51
13. MÉTODOS PARA EL AISLAMIENTO Y CARACTERIZACIÓN DE PROTEÍNAS QUE INTERACCIONAN CON LOS EXTREMOS DEL VIRUS.....	52
13.1. Fraccionamiento subcelular y preparación de extractos de proteínas.....	52
13.2. Síntesis y purificación de RNA biotinado.....	53
13.3. Cromatografía de afinidad de RNA.....	53
13.4. Preparación de muestras para su análisis mediante espectrometría de masas	54
13.5. Espectrometría de masas de desorción/ionización por laser asistida por matriz y acoplada a un detector de tiempo de vuelo (MALDI-TOF).....	55
13.6. Espectrometría de masas de ionización por electrospray acoplada a un analizador de masas de tipo cuádrupolo y un detector de tiempo de vuelo (ESI-QUAD-TOF).....	55
13.7. Herramientas informáticas para la identificación de proteínas.....	56
14. MÉTODOS PARA EL ANÁLISIS DE LA INTERACCIÓN ENTRE LOS EXTREMOS DEL VIRUS.....	56
14.1. Síntesis y purificación de RNA marcado con digoxigenina o biotina.....	56
14.2. Ensayo de coprecipitación de RNA mediada por proteínas.....	57
14.3. Electroforesis, transferencia, fijación y detección del RNA en membranas de nailon.....	58
VI. RESULTADOS.....	59
1. EFECTO DIFERENCIAL DE UNA MUTACIÓN PUNTUAL EN EL GEN DE LA REPLICASA SOBRE LA REPLICACIÓN DEL VIRUS TGEV Y LA AMPLIFICACIÓN DE MINIGENOMAS.....	59

1.1. REVERSIÓN DE UN MARCADOR GENÉTICO EN EL GENOMA DEL VIRUS RECOMBINANTE RESCATADO A PARTIR DEL CLON INFECTIVO DEL TGEV.....	59
1.2. UNA MUTACIÓN PUNTUAL EN EL GEN DE LA REPLICASA AFECTA DRÁSTICAMENTE AL RESCATE DE VIRUS A PARTIR DEL CLON INFECTIVO DEL TGEV.....	60
1.3. ESTUDIO DE LA REPLICACIÓN DE MINIGENOMAS DERIVADOS DEL TGEV.....	62
1.4. EFECTO DE UNA MUTACIÓN PUNTUAL EN EL GEN DE LA REPLICASA SOBRE EL RESCATE DEL MINIGENOMA M33L.....	67
1.5. MUTAGÉNESIS Y ANÁLISIS DEL RESCATE DEL RNA DEFECTIVO DI-C.	68
1.6. EFECTO DE UNA MUTACIÓN PUNTUAL EN EL GEN DE LA REPLICASA SOBRE PREDICCIONES DE ESTRUCTURA SECUNDARIA DEL RNA.....	71
1.7. EFECTO DE UNA MUTACIÓN PUNTUAL SOBRE EL PROCESAMIENTO <i>IN VITRO</i> NSP1/NSP2 DE LAS POLIPROTEÍNAS DE LA REPLICASA POR LA PROTEINASA VIRAL PLP1.....	73
2. GENERACIÓN Y CARACTERIZACIÓN DE UNA COLECCIÓN DE REPLICONES DERIVADOS DEL GENOMA DEL TGEV.....	77
2.1. GENERACIÓN DE UN cDNA QUE CONTIENE LAS SECUENCIAS MÍNIMAS TEÓRICAS PARA SU REPLICACIÓN AUTÓNOMA.....	77
2.2. OPTIMIZACIÓN DE LA TRANSFECCIÓN DE DNAs DE ALTA MASA MOLECULAR EN CÉLULAS EUCARIOTAS.....	79
2.3. EFECTO DEL AUMENTO DE LA EFICIENCIA DE TRANSFECCIÓN SOBRE EL RESCATE DE VIRUS A PARTIR DEL CLON INFECTIVO DEL TGEV.....	82
2.4. GENERACIÓN DE REPLICONES FUNCIONALES DERIVADOS DEL TGEV	83
2.5. ESTUDIO DE LA ACTIVIDAD DE LOS REPLICONES EN DIFERENTES LÍNEAS CELULARES.....	85
2.6. LOS REPLICONES DERIVADOS DEL TGEV NO SON CITOPÁTICOS EN LAS CÉLULAS HUMANAS 293T.....	87
2.7. EXPRESIÓN DE GENES HETERÓLOGOS A PARTIR DE UN REPLICÓN DEL TGEV.....	88

3. ESTUDIO DEL PAPEL DE LA NUCLEOPROTEÍNA VIRAL EN LA SÍNTESIS DE RNA DE CORONAVIRUS.....	90
3.1. EFECTO DE LA PROTEÍNA N SOBRE LA ACTIVIDAD DE LOS REPLICONES	90
3.2. EFECTO DE LA PROTEÍNA N O SU mRNA SOBRE LA ACTIVIDAD DE LOS REPLICONES.....	92
4. IDENTIFICACIÓN DE PROTEÍNAS VIRALES Y CELULARES QUE INTERACCIONAN CON LOS EXTREMOS 5' Y 3' DEL GENOMA	95
4.1. GENERACIÓN DE RNAs BIOTINADOS DE LOS EXTREMOS DEL GENOMA VIRAL.....	96
4.2. CROMATOGRFÍA DE AFINIDAD PARA LA CAPTURA DE PROTEÍNAS QUE INTERACCIONAN CON EL RNA VIRAL.....	98
4.3. IDENTIFICACIÓN DE PROTEÍNAS DE INTERACCIÓN CON LOS EXTREMOS DEL GENOMA.....	101
5. ESTUDIO DE LA INTERACCIÓN ENTRE LOS EXTREMOS DEL GENOMA VIRAL.....	107
5.1. DISEÑO DE UN ENSAYO DE COPRECIPITACIÓN DE RNA <i>IN VITRO</i>	107
5.2. ESTUDIO DE LOS REQUERIMIENTOS PARA LA INTERACCIÓN <i>IN VITRO</i> DE LOS EXTREMOS DEL GENOMA.....	109
5.3. ESTUDIO DE LA ESPECIFICIDAD DE LA INTERACCIÓN <i>IN VITRO</i> DE LOS EXTREMOS DEL GENOMA.....	110
VII. DISCUSIÓN.....	113
1. UN CAMBIO DE AMINOÁCIDO EN LAS POLIPROTEÍNAS DE LA REPLICASA AFECTA DE FORMA DIFERENCIAL A LA REPLICACIÓN DEL VIRUS TGEV Y A LA AMPLIFICACIÓN DE MINIGENOMAS.....	113
2. LA NUCLEOPROTEÍNA ES ESENCIAL PARA LA SÍNTESIS EFICIENTE DE RNA EN CORONAVIRUS.....	119
3. INTERACCIÓN DE PROTEÍNAS VIRALES Y CELULARES CON LOS EXTREMOS 5' Y 3' DEL GENOMA DEL TGEV.....	124
VIII. CONCLUSIONES.....	137

IX. BIBLIOGRAFÍA.....	139
X. ANEXO.....	157

ÍNDICE DE FIGURAS

1. Estructura del virión del coronavirus TGEV.....	7
2. Expresión génica del coronavirus TGEV.....	9
3. Ciclo infectivo de coronavirus.....	11
4. Sistemas de expresión basados en el genoma de coronavirus	14
5. Síntesis de macromoléculas en coronavirus.....	16
6. Modelo de trabajo de la transcripción de coronavirus.....	17
7. Organización de la replicasa de coronavirus.....	22
8. Mutagénesis puntual del genoma del TGEV	34
9. Estructura genética de los replicones derivados del TGEV.....	39
10. Estructura genética del TGEV y de RNAs defectivos derivados del mismo.....	59
11. Análisis de los marcadores genéticos en los virus rTGEV rescatados a partir del cDNA.....	60
12. Fenotipo de los virus recombinantes con mutaciones puntuales en la posición 637.....	61
13. Infectividad y producción viral en células ST y BHK-pAPN.....	63
14. Rescate de minigenomas en células BHK-pAPN y ST.....	64
15. Título de virus durante el rescate de minigenomas.....	67
16. Rescate de mutantes del minigenoma M33L.....	69
17. Rescate de mutantes del minigenoma DI-C.....	70
18. Efecto de las mutaciones en la posición 637 sobre predicciones de estructura secundaria del RNA del TGEV	72
19. Efecto de mutaciones puntuales en el nt 637 sobre el procesamiento N-terminal de la replicasa.....	74
20. Ensayo de la actividad del replicón mínimo teórico TGEV-REP.....	78
21. Esquema de la construcción pBAC-TGEV-NR-EGFP utilizada para optimizar la transfección de DNA de alta masa molecular.....	79
22. Comparación de reactivos para la transfección de BACs en células de mamífero.....	80
23. Optimización de la transfección de BAC en células 293T.....	81
24. Mejora de la eficiencia de rescate de virus infectivo como consecuencia de la optimización de la transfección de los BACs.....	83
25. Replicones derivados del TGEV mediante delección sucesiva de genes.....	84
26. Actividad de los replicones del TGEV en células humanas 293T y HeLa, células porcinas ST y células de hámster BHK-pAPN.....	86
27. Actividad de replicones del TGEV a lo largo de los pases en cultivo.....	87
28. Expresión de genes heterólogos desde el replicón REP 2 del TGEV.....	88
29. Expresión de las proteínas N y E del TGEV por el replicón del virus Sindbis en células BHK-pAPN.....	91

30. Efecto de la proteína N sobre la actividad de los replicones derivados del TGEV.....	92
31. Efecto de la proteína N y su mRNA sobre la actividad de los replicones derivados del TGEV.....	93
32. Estrategia para la identificación de proteínas que interaccionan con los extremos del genoma del TGEV.....	95
33. Generación de RNAs biotinados de los extremos 5' y 3' del genoma del TGEV	97
34. Reducción de fondo inespecífico de la cromatografía de afinidad de RNA.....	99
35. Análisis de la cantidad de RNA biotinado utilizada para el aislamiento de proteínas de unión a los extremos del genoma.....	101
36. Aislamiento e identificación de proteínas que interaccionan con los extremos 5' y 3' del TGEV en extractos celulares de tres especies distintas.....	103
37. Clasificación funcional de las proteínas identificadas por su unión a los extremos del genoma del TGEV.....	106
38. Coprecipitación de RNA de los extremos 5' y 3' del genoma en presencia de proteínas.....	108
39. Requerimientos para la coprecipitación de los extremos 5' y 3' del genoma del TGEV.....	109
40. Especificidad de la interacción <i>in vitro</i> de los extremos 5' y 3' del genoma del TGEV.....	110
41. Proteínas de fusión codificadas por los minigenomas M33L y DI-C.....	118
42. Modelo de trabajo de la replicación de coronavirus.....	134

ÍNDICE DE TABLAS

I. Grupos y especies del género coronavirus.....	5
II. Señales de replicación en los extremos del genoma de coronavirus.....	19
III. Oligonucleótidos utilizados para la mutagénesis de la posición 637 del genoma del TGEV en el plásmido pBAC-TGEV(<i>SrfI-NheI</i>).....	35
IV. Oligonucleótidos utilizados para el clonaje de secuencias de la replicasa.....	36
V. Oligonucleótidos utilizados para la generación de replicones del TGEV.....	38
VI. Oligonucleótidos utilizados para la mutagénesis de gen N.....	41
VII. Oligonucleótidos utilizados para la generación de plásmidos que contienen los extremos 5' y 3' del genoma.....	42
VIII. Oligonucleótidos utilizados para el análisis del RNA mediante RT-PCR y Q-RT-PCR.....	45
IX. Proteínas que interaccionan con los extremos 5' y 3' del genoma del TGEV.....	104
X. Proteínas que interaccionan preferentemente con el extremo 3' del genoma del TGEV.....	105

II. RESUMEN EN INGLÉS

Coronaviruses are important vertebrate pathogens with the largest RNA genomes and a replicase of enormous complexity. Coronavirus replication mechanism remains unknown, and is being studied in this thesis using the transmissible gastroenteritis coronavirus (TGEV) as a model.

The replication of full-length TGEV genome and of TGEV-derived minigenomes has been studied by reverse genetics approaches. The assembly of coronavirus replication complex requires the extensive processing of the viral polyproteins pp1a and pp1ab, leading to the synthesis of 16 mature non-structural proteins (nsps). The early proteolytic cleavage of the nsp1/nsp2 site leads to the release of nsp1. We have shown a differential requirement of this processing event for full-length genome or coronavirus-derived minigenomes replication. One amino acid substitution at position 108, that prevents proteolytic cleavage at the nsp1/nsp2 site, led to mutant viruses with low titers and small plaques compared to those of the wild type. In contrast, mutation of the cleavage site was required for efficient minigenome rescue. This finding implies possible differences in the mechanism of viral RNA replication in *cis* or the amplification in *trans* of minigenomes.

To study the requirement of structural and accessory genes for coronavirus replication, three TGEV replicons were generated by sequential deletion of genes encoding structural and accessory proteins located at the 3' end of the genome. Efficient activity of TGEV replicons was associated with the presence of the nucleoprotein (N) provided either in *cis* or in *trans*. TGEV replicons were functional in cells from different species, including human cells, in which the replicons showed a reduced cytopathic effect. A vector based in replication-competent propagation-deficient replicons has been constructed that led to high expression levels of a heterologous protein.

We have postulated the requirement of the interaction between the 5' and the 3' ends of the viral genome for coronavirus replication. Viral and cellular proteins binding to 5' and 3' ends of the TGEV genome have been identified by mass spectrometry. Viral proteins essential for coronavirus RNA synthesis such as the N protein and the RNA dependent polymerase have been identified, validating the approach. Around 30 cellular proteins from three different species binding to genome ends have been reproducibly identified. Among these proteins, the polyA binding protein, the hnRNP A1 and hnRNP A2/B showed a preferential binding for the 3' end of the genome. Another set of the

identified proteins plays a role in cellular RNA synthesis and processing and could be novel candidates for coronavirus-host interaction studies. Host protein complexes isolated by their binding to the ends of the TGEV genome mediated *in vitro* a 5'-3' crosstalk, supporting that genome circularization may be a step during coronavirus replication. To study the relevance of these interactions in coronavirus replication, RNA interference studies to suppress the expression of selected proteins are being performed.

III. INTRODUCCIÓN

1. REPLICACIÓN DE VIRUS RNA DE CADENA SENCILLA Y POSITIVA

Los virus RNA de polaridad positiva suponen un tercio de todos los géneros de virus e incluyen numerosos patógenos humanos tales como el coronavirus causante del síndrome respiratorio agudo y grave (SARS) o el virus de la hepatitis C, para los que todavía no existe una vacuna efectiva. El estudio de las bases moleculares de su replicación facilitará el desarrollo de nuevos antivirales y vacunas, que permitan combatir y prevenir las enfermedades causadas por estos virus.

La replicación de los virus RNA de cadena sencilla y positiva presenta características comunes, entre las que se encuentran: (a) el genoma viral es un RNA infectivo, que actúa de mensajero y es traducido al inicio de la infección para dar lugar a la replicasa viral; (b) la replicasa incluye una actividad RNA polimerasa dependiente de RNA (RdRp) junto con otros factores necesarios para el ensamblaje y funcionamiento de los complejos de replicación; (c) los complejos de replicación se ensamblan en membranas intracelulares, que los virus modifican a partir de distintos orgánulos; (d) la síntesis del RNA de cadena positiva es asimétrica y tiene lugar a partir de intermediarios de replicación de cadena negativa; (e) es necesario coordinar el uso del genoma infectivo como molde de traducción y síntesis de RNA; (f) debido a su capacidad codificante limitada, los virus RNA utilizan factores de la célula hospedadora para llevar a cabo estas funciones.

2. CARACTERÍSTICAS GENERALES DE LOS CORONAVIRUS

La familia *Coronaviridae* está incluida en el orden *Nidovirales*, junto con las familias *Arteriviridae* y *Roniviridae* (Cavanagh y col., 1997; Enjuanes y col., 2000b). La familia *Arteriviridae* incluye el único género *Arterivirus*, cuyo prototipo es el virus de la arteritis equina (EAV) que causa problemas reproductivos y respiratorios en caballos (Snijder y Spaan, 1995). La familia *Roniviridae* incluye el género *Okavirus*, que son los primeros *Nidovirales* que infectan a invertebrados (Walker y col., 2005). La familia *Coronaviridae* se compone a su vez de los géneros *Coronavirus* y *Torovirus*, que han sido propuestos para adquirir la categoría de familia o subfamilia (González y col., 2003) dentro del orden *Nidovirales*.

Los coronavirus (CoV) son virus esféricos con una envuelta lipídica, donde se insertan las proteínas estructurales de superficie. La proteína S, que forma las espículas, confiere a la partícula viral cierta apariencia de corona que da nombre a la familia.

El genoma de los coronavirus es un RNA no segmentado, de cadena sencilla y polaridad positiva de unas 30 kb, siendo el de mayor longitud conocida para virus RNA. La organización básica del genoma de coronavirus es común con otros miembros del orden *Nidovirales*. Todos ellos codifican las proteínas de la replicasa en los dos primeros tercios del genoma y las proteínas estructurales en el tercio próximo al extremo 3', en el orden 5'-S-E-M-N-3' (de Vries y col., 1997; Enjuanes y col., 2000a). Además, los diferentes coronavirus poseen un número variable de genes accesorios, intercalados entre los genes estructurales comunes, que codifican proteínas específicas de grupo. La expresión génica en coronavirus combina distintos mecanismos de regulación, incluyendo un salto de fase del ribosoma durante la traducción del genoma, el procesamiento proteolítico de la replicasa y la síntesis de RNAs subgenómicos 5' y 3' coterminales que se originan por un proceso característico de transcripción discontinua (Enjuanes y col., 2006; Lai y Cavanagh, 1997). La presencia de estos RNAs subgenómicos con estructura anidada dio lugar al nombre del orden *Nidovirales* (del latín *nidus*, nido).

Las especies del género *Coronavirus* se han clasificado en tres grupos en base a su homología genética y antigénica (Tabla I) (Enjuanes y col., 2000b) y constituyen patógenos que infectan a un amplio rango de vertebrados, incluyendo especies salvajes, ganaderas y animales de compañía. Los coronavirus causan mayoritariamente enfermedades respiratorias, entéricas, hepáticas y del sistema nervioso central, ocasionando graves pérdidas económicas (Lai y Holmes, 2001; Perlman y col., 2000). A finales de los años 60 los coronavirus fueron reconocidos por primera vez como patógenos humanos, causando principalmente infecciones del tracto respiratorio superior, como el resfriado común. En el año 2003, los coronavirus atrajeron la atención mundial, debido a la emergencia en China del coronavirus del SARS (SARS-CoV), que infectó a más de 8000 personas de 32 países con una mortalidad del 10% (Drosten y col., 2003; Rota y col., 2003). Desde entonces, se han identificado nuevos coronavirus en humanos, como el HCoV-NL63 (van der Hoek y col., 2004) y el HCoV-HKU1 (Woo y col., 2005), que producen neumonías relativamente graves.

TABLA I
GRUPOS Y ESPECIES DEL GÉNERO CORONAVIRUS

Grupo	Nombre	Designación ^a	Hospedador	Enfermedad
1	Virus de la gastroenteritis porcina transmisible	TGEV	Cerdo	Enteritis
	Coronavirus respiratorio porcino	PRCoV	Cerdo	Respiratoria
	Coronavirus felino	FCoV	Gato	Peritonitis
	Virus de la peritonitis infecciosa felina	FIPV	Gato	Peritonitis
	Coronavirus canino	CCoV	Perro	Enteritis
	Coronavirus humano 229E	HCoV-229E	Hombre	Respiratoria
	Virus de la diarrea epidémica porcina	PEDV	Cerdo	Enteritis
	Coronavirus humano NL63	HCoV-NL63	Hombre	Neumonía
	Coronavirus 61 del murciélago	Bat-CoV-61	Murciélago	Asintomático
	Coronavirus HKU2 del murciélago	Bat-CoV-HKU2	Murciélago	Asintomático
2	Virus de la hepatitis del ratón	MHV	Ratón	Hepatitis
	Coronavirus humano OC43	HCoV-OC43	Hombre	Enteritis
	Coronavirus del SARS	SARS-CoV	Hombre	Neumonía
	Coronavirus bovino	BCoV	Vaca	Enteritis
	Virus de la encefalitis hemaglutinante porcina	HEV	Cerdo	Encefalitis
	Coronavirus de la rata	RCoV	Rata	Respiratoria
	Virus de la sialodacrioadenitis	SDVA	Rata	Adenitis
	Coronavirus equino	ECoV	Caballo	Enteritis
	Coronavirus respiratorio canino	CRCoV	Perro	Respiratoria
	Coronavirus humano HKU1	HCoV-HKU1	Hombre	Neumonía
	Coronavirus del murciélago similar al SARS	Bat-SARS-CoV	Murciélago	Asintomático
3	Virus de la bronquitis infecciosa	IBV	Pollo	Respiratoria
	Coronavirus del pavo	TCoV	Pavo	Nefritis
	Coronavirus del faisán	PhCoV	Faisán	Respiratoria
	Coronavirus del ganso	GCoV	Ganso	Bajo peso
	Coronavirus del pato	DCoV	Pato	Bajo peso

^aEn esta memoria se utiliza la nomenclatura del género *Coronavirus* del séptimo informe del Comité Internacional de Taxonomía de Virus (ICTV) en la que al acrónimo “C” referido a coronavirus se ha sustituido por “Co” (Enjuanes y col., 2000a).

3. EL VIRUS DE LA GASTROENTERITIS PORCINA TRANSMISIBLE

El virus de la gastroenteritis porcina transmisible (TGEV), modelo de estudio en esta tesis, pertenece al grupo 1 y produce una enteritis muy contagiosa en lechones recién nacidos, con índices de mortalidad próximos al 100% (Saif y Wesley, 1992). La enfermedad, descrita por primera vez en EEUU (Doyle y Hutchings, 1946), se ha identificado en numerosos países europeos, en América central y del sur, en Canadá y

en diversos países asiáticos. Actualmente la mayor parte del ganado porcino europeo presenta inmunidad parcial frente al virus TGEV (Enjuanes y Van der Zeijst, 1995; Pensaert y Van Reeth, 1998) debido a la diseminación de un mutante natural atenuado, el coronavirus respiratorio porcino (PRCoV), mientras que en EEUU la diseminación del virus PRCoV no parece haber sido tan efectiva como en Europa (Woods y Wesley, 1998) y la enfermedad continúa siendo un problema económico para los productores de ganado porcino, particularmente en EEUU, Sudamérica, Rusia y China.

3.1. Estructura y composición del virión

El TGEV es un virus esférico con un diámetro aproximado de 100 nm, que posee una envuelta lipídica en la que se encuentran integradas las proteínas estructurales S, E y M, de 200, 28 y 10 kDa, respectivamente (Fig. 1). Por debajo de este nivel estructural se encuentra la cápsida, de apariencia densa en micrografías electrónicas de viriones maduros, que está formada por la nucleocápsida asociada al extremo C-terminal de la proteína M. La nucleocápsida helicoidal constituye el tercer nivel estructural y está compuesta por el genoma RNA asociado a la nucleoproteína N (Escors y col., 2001).

La proteína S es la más prominente de la envuelta del virus y se asocia formando trímeros que constituyen las espículas del virus (Delmas y Laude, 1990). La proteína S está glicosilada e interviene en la unión con el receptor celular y en la fusión de las membranas del virus y la célula hospedadora. Esta proteína condiciona la especificidad de tejido y de especie (Sánchez y col., 1992; Sánchez y col., 1999), y es la principal inductora de anticuerpos neutralizantes (Delmas y col., 1986; Jiménez y col., 1986; Sánchez y col., 1990).

La glicoproteína de membrana M es el componente estructural más abundante de los coronavirus (Sturman y Holmes., 1977; Sturman y col., 1980). Está formada por un segmento pequeño N-terminal situado hacia el exterior de la partícula viral, tres segmentos transmembrana y un extremo C-terminal orientado hacia el interior del virión y que interacciona con la nucleoproteína. En el caso del coronavirus TGEV, la proteína M presenta dos topologías, la segunda de ellas con ambos extremos orientados hacia el exterior de la partícula viral (Escors y col., 2001).

La proteína E es un polipéptido estructural minoritario. Esta proteína es integral de membrana, aunque todavía existe cierta ambigüedad en la descripción de su topología. Su secuencia es extremadamente divergente entre los distintos coronavirus y recientemente se ha descrito su capacidad de formar canales iónicos (Wilson y col., 2006; Wilson y col., 2004). La proteína E es esencial en el coronavirus TGEV (Ortego y col., 2002) e interviene en procesos de morfogénesis de la partícula viral.

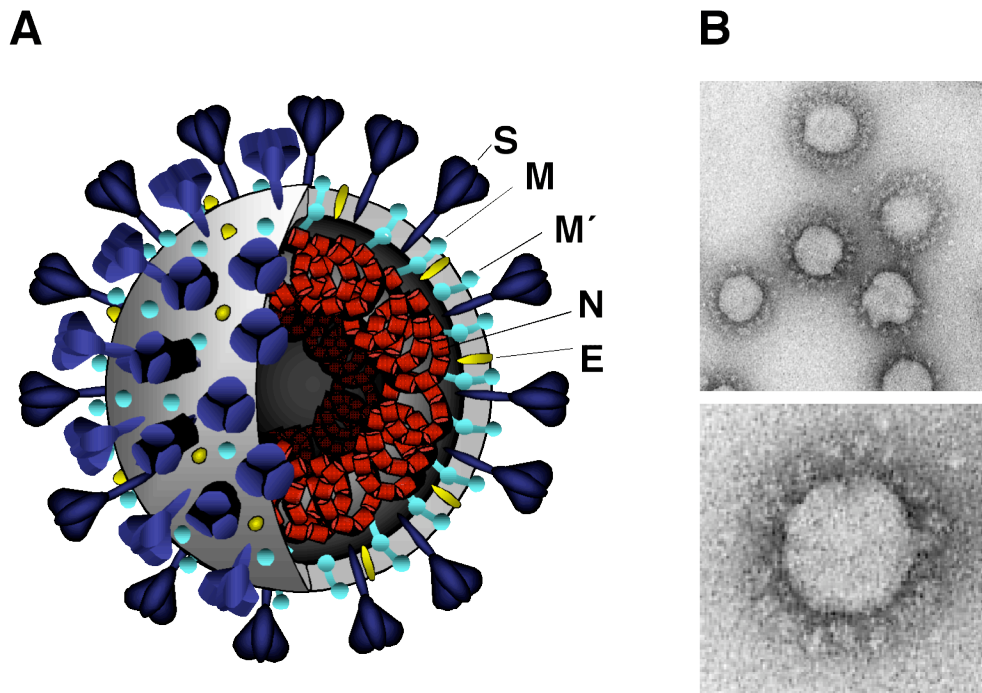


Figura 1. Estructura del virión del coronavirus TGEV. (A) Modelo de la partícula viral, con los tres niveles estructurales que la conforman: la envuelta lipídica, en la que se integran las proteínas S, E y M (en dos conformaciones M y M'), la cápsida interna constituida por la nucleoproteína asociada al endodominio C-terminal de la proteína M y la nucleocápsida helicoidal formada por el genoma RNA asociado a la proteína N. (B) Micrografías electrónicas de viriones purificados del TGEV teñidos con acetato de uranilo, mostrando las coronas características de proteínas de la espícula.

La nucleoproteína N forma parte de la nucleocápsida helicoidal mediante su asociación con el genoma viral. Recientemente, se ha propuesto una organización modular para la nucleoproteína N del SARS, basada en las regiones desordenadas de la proteína comunes entre los distintos coronavirus (Chang y col., 2006; Zúñiga y col., 2007). La proteína N tiene la capacidad de unir el RNA viral de forma específica e inespecífica, así como de formar oligómeros, lo que presumiblemente confiere estabilidad y compactación a la nucleocápsida. En la proteína N de distintos coronavirus se han localizado diferentes dominios de interacción con RNA (Fan y col., 2005; Huang

y col., 2004; Luo y col., 2006; Masters, 1992; Nelson y col., 2000) y de oligomerización (Fan y col., 2005; Hurst y col., 2005; Luo y col., 2006; Yu y col., 2005). La proteína N se une de forma específica a la secuencia reguladora de la transcripción de polaridad positiva próxima al líder (Chen y col., 2005), secuencias de la región 3' no traducida (3'UTR) (Zhou y col., 1996), al gen N (Cologna y col., 2000), y a la señal de encapsidación del genoma (Cologna y col., 2000; Molenkamp y Spaan, 1997). La proteína N está fosforilada, lo que ha sido demostrado para coronavirus de los tres grupos (Calvo y col., 2005; King y Brian, 1982; Lomniczi y Morser, 1981; Stohlman y Lai, 1979; Zakhartchouk y col., 2005) aunque el papel de esta modificación se desconoce actualmente. Se han atribuido múltiples funciones a la proteína N, además de formar parte estructural del virión. Sin embargo, muchas de estas funciones no se han demostrado formalmente. Debido a la interacción de la proteína con secuencias reguladoras de la transcripción se le ha asignado un posible papel en la transcripción viral, y se ha propuesto que podría intervenir en la traducción preferente de los RNAs virales frente a los celulares, por su unión a la secuencia líder del extremo 5' del genoma (Tahara y col., 1998; Tahara y col., 1994). Una población de la proteína N ha sido localizada en el nucleolo de células infectadas por coronavirus de los tres grupos (Hiscox y col., 2001; Wurm y col., 2001), donde colocaliza con proteínas nucleolares como la fibrilarina, e interacciona con la nucleolina (Chen y col., 2002). Sin embargo, la localización nucleolar de la proteína N probablemente dependa del tipo celular donde se crece el virus, dado que no ha sido confirmada por otros autores para los coronavirus TGEV (Calvo y col., 2005) o SARS-CoV (Rowland y col., 2005).

3.2 Estructura del genoma y expresión génica

El genoma del TGEV es un RNA no segmentado, de cadena sencilla y polaridad positiva de 28.5 kb de longitud (Fig. 2), que es infeccioso una vez introducido en células susceptibles (Norman y col., 1968). Este genoma, se asemeja estructuralmente en sus extremos a los mensajeros celulares por tener una estructura de 7-metil-guanosina o cap en el extremo 5' (Page y col., 1990), estar poliadenilado en el extremo 3' (Jacobs y col., 1986) y poseer regiones no traducidas en ambos extremos (5'UTR y 3'UTR). Sin embargo, se diferencia notablemente de los mensajeros eucariotas en su tamaño extremadamente grande y en su estructura policistrónica. El genoma está formado por una secuencia líder de 99 nt en el extremo 5', la región 5'UTR, el gen de la

replicasa que ocupa unas 20 kb, y los genes S, 3a, 3b, E, M, N y 7, seguidos de la región 3'UTR. En algunos casos existe solapamiento parcial de secuencia entre genes adyacentes. El gen de la replicasa es traducido al inicio de la infección utilizando dos ORFs parcialmente solapantes, ORF1a y ORF1b, que dan lugar a dos poliproteínas co-amino terminales, pp1a y pp1ab, la segunda de las cuales se traduce mediante un salto de fase del ribosoma (Brierley y col., 1989). Este mecanismo, depende de dos elementos de RNA, un heptanucleótido que actúa como secuencia de deslizamiento del ribosoma seguido de un falso lazo o *pseudoknot* que dificulta el progreso de la traducción (Somogyi y col., 1993). Con una frecuencia en torno al 20-30%, estas estructuras ocasionan el retroceso de una base del ribosoma y el cambio de marco de lectura que da lugar a la poliproteína pp1ab (Baranov y col., 2005; Eleouet y col., 1995).

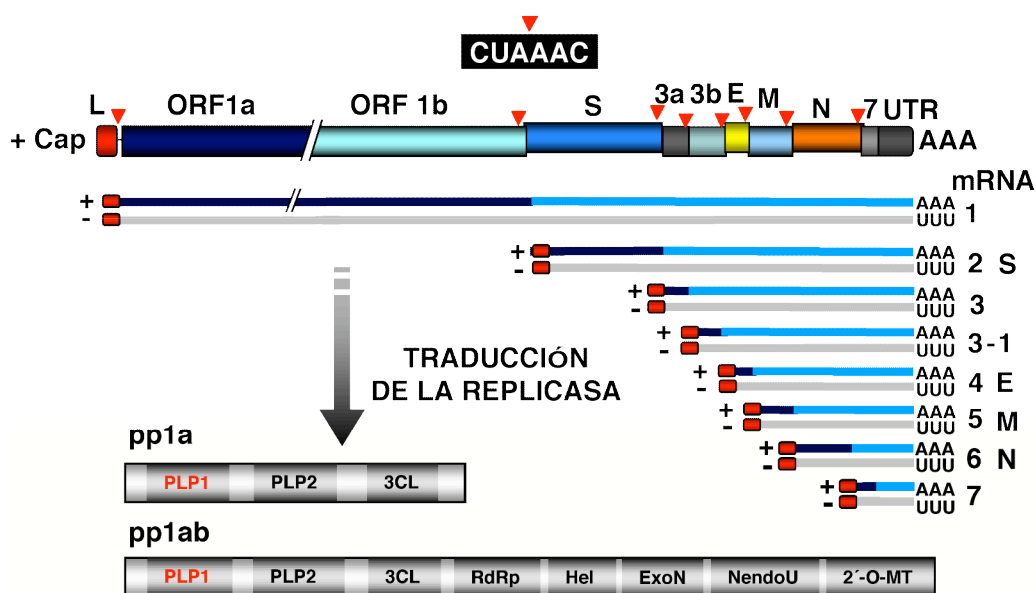


Figura 2. Expresión génica del coronavirus TGEV. Esquema del genoma del TGEV (barra superior), en el que las letras y números indican los genes virales. L, secuencia líder; UTR, región 3' no traducida; AAA, poliA. Los distintos RNAs virales, genómico y subgenómicos, están numerados del 1 al 7 y se indica su polaridad y estructuras terminales comunes. Las secuencias líder y antilíder se muestran en rojo y la secuencia codificante traducida a partir de cada mensajero policistrónico en azul oscuro. Las secuencias reguladoras de la transcripción o TRSs se indican con triángulos rojos invertidos. En la parte superior de la figura se muestra la secuencia conservada o CS. Las poliproteínas de la replicasa, sin procesar, se representan por barras grises en la parte inferior, con las principales actividades enzimáticas que contienen. PLP1 y PLP2, proteasas similares a papaína 1 y 2; 3CL, proteasa principal; RdRp, RNA polimerasa dependiente de RNA, Hel, helicasa; ExoN, exoribonucleasa; NendoU, endoribonucleasa; 2'-O-MT, ribosa 2'-O-metiltransferasa.

Las poliproteínas pp1a y pp1ab se autoprocen para liberar los componentes del complejo de replicación-transcripción. A diferencia del gen de la replicasa, los genes situados en el extremo 3' del genoma se expresan a partir de una colección de

mensajeros subgenómicos 5' y 3' coterminales (Fig. 2). Todos ellos presentan en el extremo 5' la secuencia líder, la cual se incorpora mediante un proceso de síntesis discontinua que es asistido por las secuencias reguladoras de la transcripción o TRSs. Estas secuencias se localizan en el extremo 3' de la secuencia líder y precediendo a cada uno de los genes y, en el caso del TGEV, están formadas por el hexanucleótido central conservado (CS) 5'-CUAAAC-3' flanqueado por secuencias de longitud variable denominadas 5' y 3' TRSs que son importantes para la funcionalidad de la CS (Sola y col., 2005). Los mensajeros subgenómicos resultantes son estructuralmente policistronicos, pero por lo general sólo se traduce en cada caso la ORF situada en su extremo 5'-proximal.

3.3. Ciclo infectivo viral

La infección por coronavirus (Fig. 3) comienza con la unión de la proteína S de la superficie de los viriones con el receptor celular, que en el caso del TGEV es la aminopeptidasa N porcina (pAPN) (Delmas y col., 1992), lo que induce la entrada del virus. Una vez liberada la nucleocápsida en el citoplasma de la célula, el genoma RNA es traducido, dando lugar a las poliproteínas pp1a y pp1ab que se autoprocenan por tres proteasas virales dando lugar a los componentes virales del complejo de replicación-transcripción. Este complejo se ensambla en vesículas de doble membrana (Gosert y col., 2002; Snijder y col., 2006) y contiene presumiblemente factores celulares. La replicasa utiliza el genoma entrante, de polaridad positiva, como molde para la síntesis de los genomas progenie y de los mensajeros subgenómicos, a través de intermediarios de polaridad negativa de tamaño genómico y subgenómico, respectivamente (Sawicki y Sawicki, 1990; Sawicki y col., 2007). A partir de los RNAs subgenómicos se traducen las proteínas estructurales y no estructurales del extremo 3' del genoma. Las proteínas de la envuelta S, M y E se insertan en la membrana del retículo endoplásmico a partir del que se translocan a las cisternas del compartimento intermedio. Las nucleocápsidas virales, formadas por los genomas progenie asociados a la proteína N, se asocian con los componentes estructurales integrados en las membranas del compartimento intermedio, a través de la interacción entre la nucleoproteína y el dominio C-terminal de la proteína M (Narayanan y Makino, 2001; Sturman y col., 1980). Este proceso induce una invaginación de la membrana y envoltura de la nucleocápsida que da lugar a los viriones inmaduros. Finalmente, los viriones progenie son exportados en cisternas del aparato de Golgi, donde ocurre su maduración (Salanueva y col., 1999), hasta la

membrana plasmática para ser liberados al espacio extracelular, mediante la fusión de las vesículas que los contienen con la membrana plasmática.

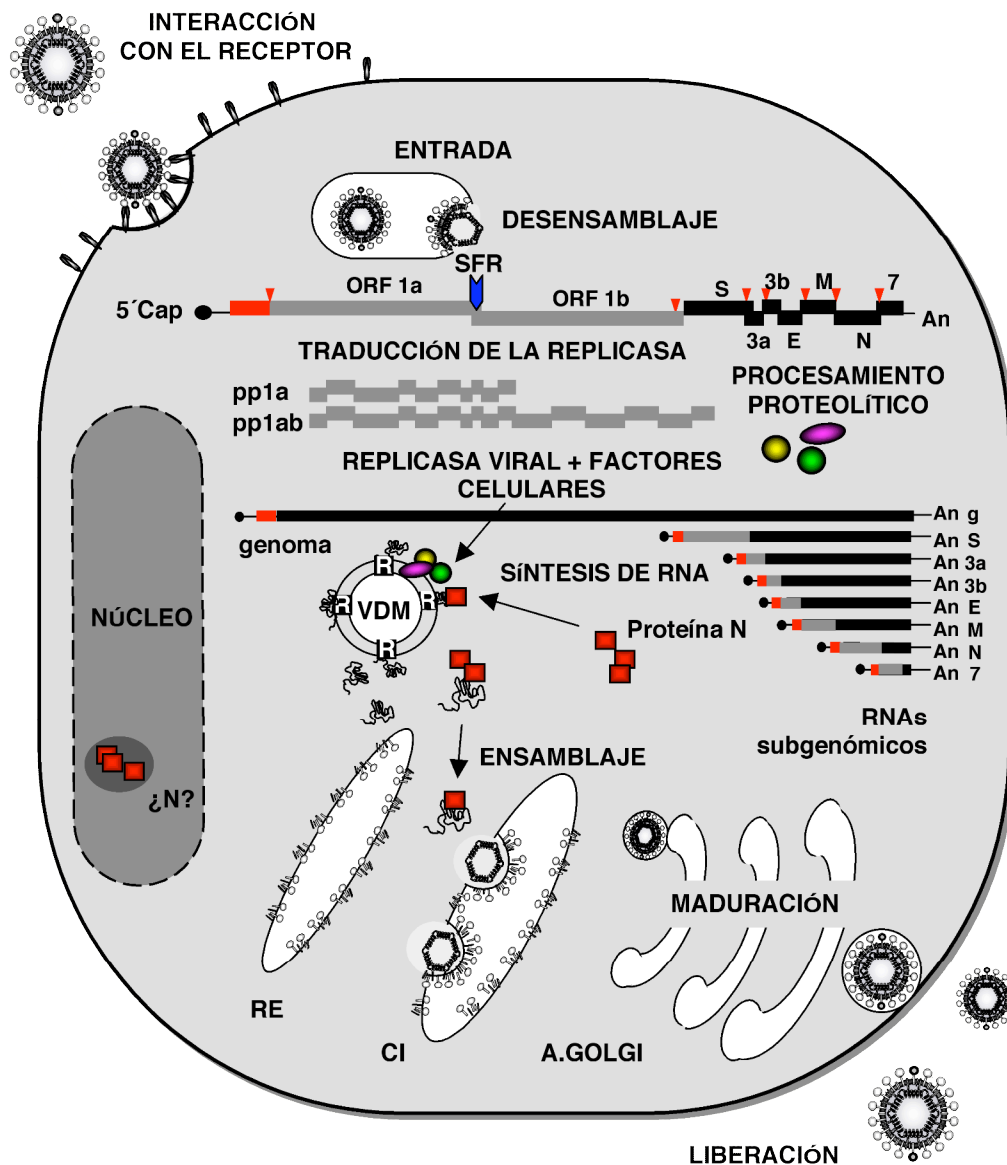


Figura 3. Ciclo infeccioso de coronavirus. Esquema de la infección por el TGEV. Por simplificación, sólo representan las especies de RNA codificantes, de polaridad positiva. SFR, salto de fase del ribosoma; An, poliA; R, complejo de replicación; VDM, vesículas de doble membrana; RE, retículo endoplásmico; CI, compartimento intermedio; A.Golgi, aparato de Golgi.

4. GENÉTICA REVERSA EN CORONAVIRUS

La manipulación del genoma de coronavirus mediante genética reversa, ha estado limitada por el enorme tamaño del genoma de los coronavirus y la existencia de regiones tóxicas, en el gen de la replicasa, que impedían el clonaje y crecimiento en bacterias de un cDNA infeccioso. Inicialmente, los estudios genéticos en coronavirus

estaban limitados al análisis de RNAs defectivos o mutantes termosensibles. El primer sistema de manipulación del genoma completo de coronavirus, fue desarrollado originalmente para el virus MHV (Masters, 1999) y estaba basado en la recombinación dirigida de RNA, aprovechando la alta tasa de recombinación homóloga de este virus. Sin embargo, este sistema no permitía manipular con facilidad el extremo 5' del genoma ni las secuencias próximas al extremo 3' del mismo, por lo que era necesario generar sistemas de genética reversa que permitieran modificar estas secuencias.

4.1. Sistemas dependientes de virus complementador

Hasta la obtención de cDNAs virales de longitud completa, los estudios de expresión génica en coronavirus y la caracterización de las señales en *cis* necesarias para su replicación se realizaron utilizando RNAs defectivos o minigenomas, naturales o sintéticos, que requieren la presencia de un virus complementador para replicarse. En el coronavirus TGEV, el RNA defectivo interferente DI-C de 9.7 kb (Izeta y col., 1999) (Fig. 10) se utilizó para generar una colección de mutantes de delección y se desarrollaron dos aproximaciones para evaluar su amplificación por el virus complementador. En primer lugar, se sintetizaron los minigenomas por transcripción *in vitro*, el RNA se introdujo por electroporación en células ST infectadas y se estudió su rescate a lo largo de los pases. El minigenoma M33, de 3.3 kb, fue el RNA sintético de menor tamaño rescatado con alta eficiencia (Izeta y col., 1999) (Fig. 10).

Posteriormente se diseñó una estrategia de amplificación de los minigenomas en dos etapas. Para ello, su cDNA se clonó bajo el control del promotor inmediatamente temprano de citomegalovirus (CMV) y se flanqueó en el extremo 3' por la secuencia de la ribozima del virus de la hepatitis delta (HDV) y la secuencia de terminación de la transcripción y poliadenilación del gen de la hormona de crecimiento bovina (BGH), para el procesamiento y terminación correcta del RNA. En este sistema, el RNA del minigenoma se transcribe en el núcleo por la RNA polimerasa II celular y es translocado al citoplasma donde tiene lugar una segunda amplificación por la RdRp del virus complementador. Este sistema permitió la expresión de genes heterólogos *in vitro* e *in vivo* (Alonso y col., 2002).

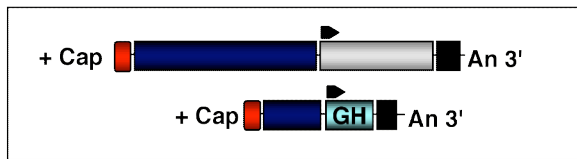
4.2. Generación de clones infectivos de coronavirus

El primer clon infectivo de coronavirus se construyó en nuestro laboratorio para el TGEV y fue ensamblado en un cromosoma artificial de bacterias (BAC) (Almazán y col., 2000). Este sistema limita la cantidad de vector por bacteria de una a dos copias, reduciendo así la toxicidad del cDNA viral durante su propagación en *Escherichia coli*. El ensamblaje se realizó tomando como base un clon cDNA del RNA defectivo natural DI-C, formado por cuatro segmentos discontinuos del genoma del TGEV, a partir del cual se reconstruyó el genoma completo. Este procedimiento permitió la localización de la región del gen de la replicasa responsable de la toxicidad del cDNA viral en las bacterias, que se introdujo en un último paso de clonaje. El cDNA del TGEV se clonó bajo el promotor de CMV y se flanqueó en su extremo 3' por las secuencias reguladoras para la terminación correcta del RNA, utilizando la misma estrategia descrita para el sistema de doble amplificación de minigemomas. Una vez transfectado el clon infectivo en células eucariotas, el genoma viral se expresa en el núcleo por la RNA pol II celular y posteriormente se transloca al citoplasma, donde comienza el ciclo infectivo del virus. Esta estrategia ha sido utilizada posteriormente con éxito para la generación de clones infectivos de los virus HCoV-OC43 (St-Jean y col., 2006) y SARS-CoV (Almazán y col., 2006).

Una estrategia independiente para obtener un clon infectivo de coronavirus fue la generación de un cDNA viral como molde de transcripción *in vitro*, reconstruido mediante la ligación *in vitro* de fragmentos de cDNA que comprenden todo el genoma (Fig. 4). Utilizando esta aproximación se han generado clones infectivos de los coronavirus TGEV (Yount y col., 2000), MHV-A59 (Yount y col., 2002) y SARS-CoV (Yount y col., 2003).

En una tercera estrategia de genética reversa, el cDNA del coronavirus humano HCoV-229E, precedido por el promotor del fago T7, se generó mediante la ligación *in vitro* de tres fragmentos de cDNA, y se insertó en el genoma del virus vaccinia, utilizado como vector de clonaje del cDNA viral (Fig. 4). El genoma del virus vaccinia purificado se utilizó posteriormente como molde para la transcripción *in vitro* del genoma del HCoV-229E (Thiel y col., 2001a). Utilizando una variante de esta estrategia se han generado clones infectivos para los coronavirus IBV (Casais y col., 2001) y MHV (Coley y col., 2005).

A. DEPENDIENTES DE VIRUS COMPLEMENTADOR



B. BASADOS EN GENOMA ÚNICO

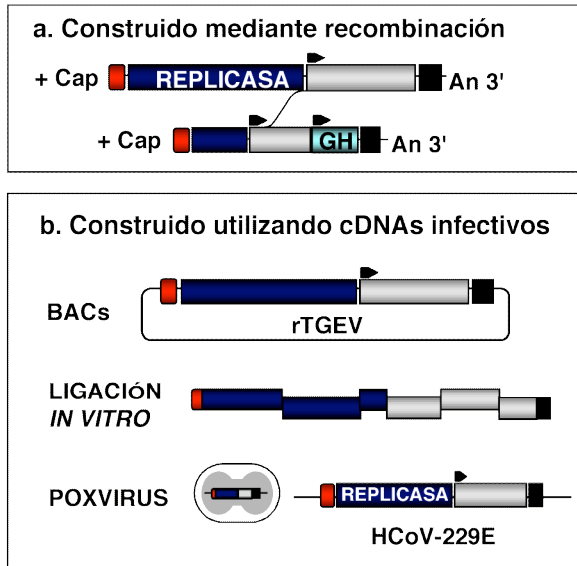


Figura 4. Sistemas de expresión basados en el genoma de coronavirus. (A) Sistema de expresión que utiliza dos componentes, basado en un virus complementador y un RNA defectivo que contiene el gen heterólogo (GH). An, poliA. (B) Sistemas basados en un genoma único. (a) Estrategia de inserción de un gen heterólogo en el genoma mediante recombinación dirigida. (b) Estrategias para la construcción de clones infectivos de coronavirus basados en el clonaje del cDNA viral en un BAC, la ligación *in vitro* de fragmentos de cDNA, o el clonaje del cDNA viral en el genoma de poxvirus.

4.3. Los coronavirus como vectores de expresión

El desarrollo de vectores de expresión basados en el genoma de virus RNA de cadena positiva ha sido utilizado con éxito para picornavirus (Andino y col., 1994), flavivirus (Chambers y col., 1999; Khromykh y Westaway, 1994) y alfavirus (Caley y col., 1999; Frolov y col., 1996; Liljestrom, 1994; Pushko y col., 1997). En la actualidad los sistemas de expresión más avanzados están basados en los alfavirus, incluyendo el virus Sindbis, el virus del bosque de Semliki y el virus de la encefalitis equina de Venezuela, los cuales son muy eficientes en desencadenar una respuesta inmune humoral y celular.

Los coronavirus presentan varias ventajas para el desarrollo de vectores: (a) son virus RNA que replican en el citoplasma de las células infectadas sin un intermediario DNA, lo que hace altamente improbable su integración en el genoma de la célula (Lai y Cavanagh, 1997); (b) poseen el genoma RNA de mayor tamaño, lo que les confiere una alta capacidad teórica de clonaje (Masters, 1999); (c) infectan principalmente a través de mucosas, por lo que pueden ser utilizados para inducir una respuesta inmune secretora (Enjuanes y col., 2003; 2005); (d) su tropismo puede ser modificado mediante

la manipulación de la proteína S (Ballesteros y col., 1997; Kuo y col., 2000; Leparco-Goffart y col., 1997; Sánchez y col., 1999); (e) existen cepas no patogénicas que infectan la mayoría de las especies de interés, de utilidad para desarrollar vectores de expresión (Sánchez y col., 1992).

El desarrollo de vectores basados en el genoma de coronavirus requiere de estudios básicos y aplicados, con objeto de optimizar los niveles de expresión, determinar los genes no esenciales con el fin de generar espacio para la inserción de genes heterólogos, y diseñar estrategias para garantizar la bioseguridad de los vectores (Enjuanes y col., 2003; 2005).

En coronavirus se han desarrollado dos tipos de sistemas para la expresión de genes heterólogos, uno basado en dos componentes, que es dependiente de un virus complementador y otros basados en el uso de genomas únicos (Fig. 4). Inicialmente se utilizó el sistema de dos componentes para la expresión de genes heterólogos en MHV (Lin y Lai, 1993), y posteriormente se han desarrollado sistemas análogos en coronavirus de los tres grupos. Utilizando el minigenoma M39 de 3.9 kb derivado del TGEV, se ha expresado el gen de la β -glucuronidasa y el antígeno de superficie ORF 5 del virus del síndrome respiratorio y reproductivo porcino (PRRSV), obteniéndose altos niveles de expresión (Alonso y col., 2002). Utilizando el clon infectivo del TGEV, se ha expresado la proteína verde fluorescente (GFP) reemplazando los genes 3a y 3b, obteniéndose niveles de expresión del orden de $> 40 \mu\text{g}/10^6$ células. El virus recombinante generado indujo una respuesta inmune lactogénica frente a la proteína heteróloga en cerdas lactantes y su progenie (Sola y col., 2003), validando la eficiencia del sistema. Actualmente, la disponibilidad de clones infectivos de coronavirus, plantea la posibilidad de generar replicones, eficientes en replicación y deficientes en propagación, para estudiar las bases moleculares de la replicación y la transcripción sin necesidad de manejar virus infectivo, así como para el desarrollo de vectores virales bioseguros.

5. SÍNTESIS DE RNA EN CORONAVIRUS

La síntesis de RNA en coronavirus es un proceso complejo y en gran parte desconocido llevado a cabo por el complejo de replicación-transcripción, constituido por las proteínas codificadas en el gen de la replicasa y posiblemente por proteínas celulares. La replicación y la transcripción de los coronavirus tienen lugar en unas

vesículas citoplasmáticas de doble membrana que contienen todas las proteínas necesarias para la síntesis del RNA viral (Gosert y col., 2002; Snijder y col., 2006). Mecánicamente, la replicación de coronavirus plantea el problema de coordinar la utilización de una misma molécula molde, el genoma de polaridad positiva, para la síntesis de las poliproteínas del gen de la replicasa mediante traducción, la síntesis de los genomas progenie mediante replicación y la producción de los RNAs subgenómicos mediante transcripción (Fig. 5). Las diferencias entre estos tres procesos implican la necesidad de mecanismos de control, de tipo espacial y temporal, posiblemente regulados mediante la participación de distintas proteínas virales o celulares.

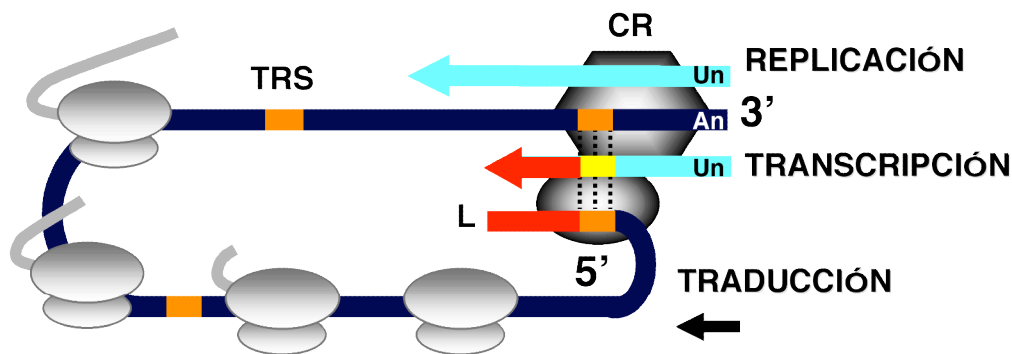


Figura 5. Síntesis de macromoléculas en coronavirus. Esquema del genoma viral de polaridad positiva (en azul oscuro), que es utilizado como molde para la traducción de la replicasa, y la síntesis del RNA genómico (replicación) ó de los mRNAs (transcripción). El sentido de la síntesis de proteínas (5'-3') y de RNA (3'-5'), se indica con flechas. TRS, secuencia reguladora de la transcripción; L, líder; CR, complejo de replicación-transcripción; An, Un, poliA y poliU, respectivamente.

5.1. Transcripción

Para explicar el mecanismo de transcripción discontinua de coronavirus, se han propuestos varios modelos durante las dos últimas décadas. Inicialmente, se demostró que los RNAs subgenómicos se producían por un mecanismo de síntesis discontinua y no mediante procesamiento o *splicing* a partir de un precursor de tamaño genómico (Jacobs y col., 1981; Stern y Sefton, 1982). Recientemente, también se ha descartado que se produzca *splicing* durante la síntesis de la cadena negativa (Sawicki y col., 2007). Un primer modelo de transcripción, denominado “modelo de transcripción iniciada por líder” (Lai, 1986; Sethna y col., 1989) proponía que la síntesis discontinua ocurría durante la generación de los mensajeros de polaridad positiva. El modelo de transcripción actual propone sin embargo que el paso de síntesis discontinua ocurre durante la transcripción de la cadena negativa, utilizando como molde el RNA

genómico de polaridad positiva (Fig. 6) [revisado en (Sawicki y Sawicki, 2005; Sawicki y col., 2007)]. Este modelo se propuso al identificar intermediarios de replicación de polaridad negativa de tamaño genómico y subgenómico en células infectadas con el virus MHV-A59 (Sawicki y Sawicki, 1990). La síntesis comenzaría copiando el extremo 3' del genoma, hasta alcanzar las distintas secuencias reguladoras de la transcripción, que, actuarían como señales de parada para la polimerasa. A continuación se produciría el apareamiento de la TRS complementaria naciente con la TRS del líder, lo que supone un cambio de molde, siguiendo con la síntesis del líder (van Marle y col., 1999; Zúñiga y col., 2004). A partir de los RNAs subgenómicos de polaridad negativa se produciría la síntesis asimétrica y continua de los mensajeros de polaridad positiva. Uno de los datos experimentales que apoyan este modelo es que, en general, la síntesis de los mensajeros subgenómicos ocurre en cantidad proporcional a su distancia al extremo 3' del genoma (Joo y Makino, 1995; Krishnan y col., 1996; Ozdarendeli y col., 2001). Este mecanismo de transcripción implica la necesidad de aproximación entre regiones de secuencia distales en el genoma, que podría estar mediada por proteínas virales o celulares o mediante interacciones RNA-proteína y proteína-proteína (Zúñiga y col., 2004).

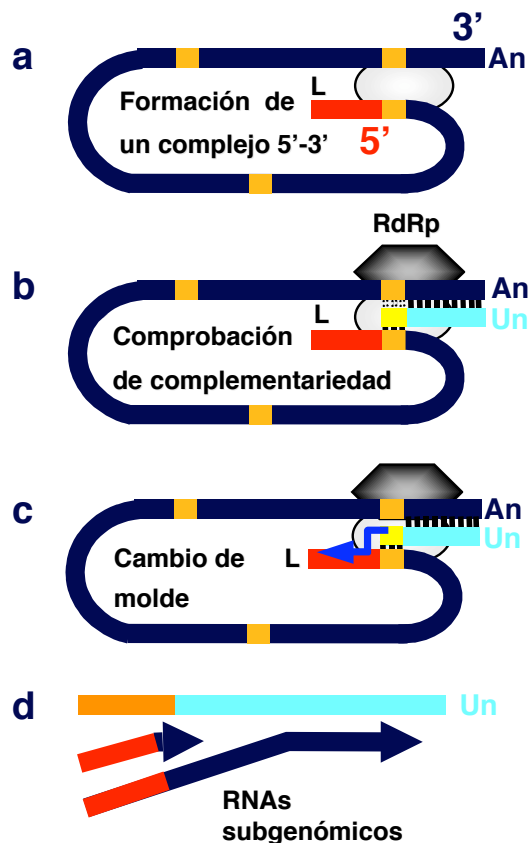


Figura 6. Modelo de trabajo de la transcripción de coronavirus. La transcripción discontinua se asemeja a una recombinación de alta frecuencia dependiente de homología y el modelo comprende varias etapas: (a) formación de un complejo entre los extremos del genoma; (b) síntesis de la cadena negativa y comprobación de la complementariedad entre las TRSs de la cadena naciente y la TRS del líder; (c) cambio de molde eventual y copia del líder; (d) síntesis asimétrica y continua de los mensajeros de polaridad positiva. Línea azul oscuro, genoma viral; línea azul claro, mensajero de polaridad negativa; línea naranja, TRS; línea amarilla, TRS complementaria; flecha azul quebrada, cambio de molde; línea roja, secuencia líder.

5.2. Replicación

A diferencia del proceso discontinuo de transcripción característico de los nidovirales, la replicación es continua y tiene lugar por un mecanismo asimétrico a partir de un intermediario de RNA de longitud completa y polaridad negativa que actúa como molde eficiente para la síntesis de nuevas copias del genoma [revisado en (Brian y Baric, 2005)]. Actualmente se desconocen las proteínas específicas que intervienen en la replicación y los mecanismos que regulan la síntesis asimétrica de las dos polaridades del RNA. Teniendo en cuenta el gran tamaño del genoma de coronavirus (28-30 kb) y la relativamente baja fidelidad de la RdRp, la replicación de coronavirus debe tener características únicas que garanticen la replicación eficiente del genoma sin que se produzca el fenómeno conocido como catástrofe de error (Gorbalenya y col., 2006).

5.2.1. Señales en *cis* necesarias para la replicación

La identificación de distintos RNAs defectivos derivados de coronavirus, capaces de ser replicados en *trans* por el virus complementador, ha permitido estudiar los elementos de secuencia en *cis* necesarios para la replicación de estos RNAs, y presumiblemente, del genoma viral. Utilizando RNAs defectivos de los virus TGEV, MHV, BCoV e IBV, se han realizado análisis de delección, y para la mayoría de estos minigenomas las únicas secuencias necesarias para su replicación residen en sus extremos 5' y 3', derivados del genoma parental. La extensión de estas secuencias varía en función del RNA defectivo estudiado, siendo de 466-1348 nt para el extremo 5' y de 338-1635 nt para el extremo 3' (Tabla II). Sin embargo, los requerimientos mínimos no se han determinado de forma precisa y estudios de delección posteriores podrían acotar estas secuencias con exactitud.

El minigenoma más pequeño derivado del TGEV que se rescató con alta eficiencia fue el M33, de 3.3 kb (Izeta y col., 1999), formado por cuatro discontinuidades respecto del genoma viral que comprenden: (a) los primeros 2144 nt y los 174 nt del extremo 3' de la ORF1a; (b) los primeros 394 nt y los 11 nt del extremo de 3' de la ORF1b; (c) los primeros 8 nt del gen S; y (d) los 492 nt del extremo 3' del genoma, incluyendo parte del gen 7, y la 3'UTR completa. Este minigenoma expresa teóricamente una ORF de fusión de 2.3 kb, que representa el 70% de la longitud del minigenoma. El minigenoma más pequeño que se pudo estudiar fue el M21, de 2.1 kb (Izeta y col., 1999), aunque este minigenoma se rescataba con una eficiencia unas 100 veces menor que el M33 (C.

Capiscol, comunicación personal) debido a un defecto de la replicación o encapsidación del RNA por el virus complementador. La amplificación del minigenoma M21 se evaluó en el pase P0 mediante la expresión desde el minigenoma del gen trazador de la β -glucuronidasa y no se observó un defecto en su replicación, por lo que se trataba probablemente de un defecto en su encapsidación. En base a estos estudios, las secuencias en *cis* necesarias para la replicación del TGEV estarían acotadas por el minigenoma M21, que contiene los primeros 1348 nt del extremo 5' y los 492 nt del extremo 3'.

Los requerimientos de secuencia para la síntesis de la cadena positiva y negativa se han estudiado utilizando el minigenoma DissE de MHV (Lin y col., 1994). Con esta construcción se determinó que los requerimientos para la síntesis de la cadena positiva eran 470 nt del extremo 5' y 460 nt del extremo 3'. Sin embargo, únicamente los últimos 55 nt del extremo 3' más la cola de poliA eran suficientes para la síntesis de la cadena negativa. Teóricamente, los promotores necesarios para la replicación de una cadena deberían residir en el extremo opuesto de la cadena complementaria. Por el contrario, estas observaciones indicaban que señales en *cis* en los dos extremos 5' y 3' del genoma eran necesarias para la síntesis de la cadena positiva, lo que sugiere una comunicación entre ambos extremos del genoma durante la replicación. En los otros coronavirus no se han estudiado los requerimientos necesarios para la síntesis de cada una de las cadenas del RNA.

**TABLA II. SECUENCIAS MÍNIMAS NECESARIAS
PARA LA REPLICACIÓN DEL GENOMA DE CORONAVIRUS**

Grupo	Virus	Extremo 5'	Extremo 3'	Referencias
1	TGEV	1348	492	(Izeta y col., 1999)
2	MHV	466	436	(Lin y Lai, 1993; Luytjes y col., 1996; Makino y col., 1990; van der Most y col., 1995; van der Most y col., 1991)
	BCoV	498	ND ^b (1637)	(Chang y col., 1994)
3	IBV	544	338	(Dalton y col., 2001)

^bND; no determinado.

A pesar de su utilidad en la acotación de las señales de replicación en *cis*, el uso de los minigenomas para estudiar la replicación tiene algunas limitaciones como son la alta frecuencia de recombinación entre estos RNAs defectivos y el virus complementador, lo que restaura a menudo las mutaciones introducidas en los minigenomas, y la necesidad de un paso previo de amplificación para detectar los minigenomas.

5.2.2. Estructuras de RNA implicadas en la replicación

Los extremos del genoma de coronavirus, donde residen las señales de replicación en *cis*, contienen superestructuras de RNA que son esenciales en la replicación de los RNAs defectivos. Todos los estudios se han llevado a cabo utilizando los minigenomas de menor tamaño derivados de los virus MHV y BCoV. En el extremo 5' del virus BCoV se han identificado cuatro horquillas (numeradas de I a IV) (Chang y col., 1996; Chang y col., 1994), que están parcialmente conservadas en el coronavirus TGEV, en base a predicciones estructurales. La horquilla III contiene la primera fase de lectura abierta del genoma, cuya traducción daría lugar a un péptido de tres aminoácidos, y comprende los nucleótidos 102-140 del TGEV por homología con el coronavirus BCoV (Raman y col., 2003). La horquilla IV precede a la segunda fase de lectura abierta, que da lugar por traducción a la replicasa, y comprende los nucleótidos 266-296 del TGEV (Raman y Brian, 2005). Estas dos estructuras han sido implicadas funcionalmente en la replicación de RNAs defectivos.

En el extremo 3' del TGEV se han identificado dos estructuras conservadas (Izeta, 2000), una horquilla que comprende los nucleótidos 274-217 desde el poliA precedida por el codón de terminación del gen N, homóloga a la de MHV (Hsue y Masters, 1997), y una estructura conservada de falso lazo o *pseudoknot* entre los nucleótidos 213-160, homóloga a la del BCoV (Williams y col., 1999). Ambos elementos han sido implicados funcionalmente en la replicación de los RNAs defectivos de MHV y BCoV, respectivamente.

6. LA REPLICASA DE CORONAVIRUS

El complejo de replicación está formado presumiblemente por los productos del gen de la replicasa y por proteínas celulares. Las poliproteínas pp1a y pp1ab, resultantes de la traducción del gen de la replicasa mediante un salto de fase del ribosoma, son procesadas por dos o tres proteasas codificadas en la ORF1a, dando lugar a un gran

número de intermediarios de procesamiento y productos maduros [revisado en (Ziebuhr, 2005)]. Este mecanismo de control doble de la expresión de la replicasa, podría regular la estequiometría de los componentes del complejo de replicación, y facilitar en primer lugar el ensamblaje de los componentes estructurales, mayoritariamente codificados en la ORF1a, y la participación posterior de las enzimas de síntesis y procesamiento del RNA, codificadas en la ORF1b (Fig. 7).

Los productos finales del procesamiento de la replicasa son 16 proteínas denominadas no estructurales (nsp), salvo en el coronavirus IBV que carece de la nsp1 (Ziebuhr, 2005). Los productos derivados de la poliproteína pp1a (nsp1-nsp11) son comunes con los procesados a partir de la pp1ab (nsp1-nsp10 y nsp12-nsp16) con la excepción de la nsp11, que es un polipéptido originado cuando no tiene lugar el salto de fase del ribosoma (Fig. 7).

Todos los coronavirus codifican una proteasa principal denominada 3CL o M^{pro} (nsp5) responsable de su autoproteolisis y del corte proteolítico de todos los dominios situados a partir de su extremo C-terminal, y dos proteasas similares a papaína (PLP1 y PLP2) que se localizan en la nsp3 y que procesan las proteínas del extremo N-proximal de la replicasa. En el caso de los virus IBV y SARS-CoV sólo poseen una proteasa similar a papaína activa, la PLP2 (Ziebuhr, 2005; Ziebuhr y col., 2000) (Fig. 7). Estas dos proteasas parecen haberse originado mediante duplicación y han mostrado cierta redundancia funcional en ensayos *in vitro* (Ziebuhr y col., 2001). La proteasa PLP1 de MHV procesa los sitios nsp1/nsp2 y nsp2/nsp3, mientras que la actividad PLP2 procesa el tercer sitio nsp3/nsp4. En el caso del coronavirus humano HCoV-229E, la proteasa PLP2 también puede procesar los sitios nsp1/nsp2 y nsp2/nsp3 además del sitio nsp3/nsp4 (Ziebuhr, 2005; Ziebuhr y col., 2000). En el virus SARS-CoV la proteasa PLP2 procesa los dos sitios N-proximales nsp1/nsp2 y nsp2/nsp3 (Han y col., 2005), y en el caso del virus IBV, la proteasa activa PLP2 procesa el sitio nsp2/nsp3 (Lim y Liu, 1998), puesto que este virus carece del sitio de procesamiento más N-terminal equivalente respecto de los otros coronavirus. Las proteínas nsp1 y nsp2 se diferencian en secuencia y tamaño entre los coronavirus del grupo 1 (proteínas p9 y p87) y del grupo 2 (proteínas p28 y p65) y su función es desconocida. La proteína nsp1 es la primera forma madura que se procesa en células infectadas (Denison y col., 1992; Herold y col., 1998). En el coronavirus MHV se ha observado que las proteínas nsp1 y nsp2 no son estrictamente necesarias para la infección, pero su eliminación o mutación

produce retraso en el crecimiento del virus, disminución en los títulos virales, o ambos (Brockway y Denison 2005; Graham y Denison, 2006; Graham y col., 2005).

La replicasa de los coronavirus contiene además actividades RNA polimerasa dependiente de RNA (nsp12) y helicasa de RNA (nsp13), que son comunes en los virus RNA. A partir de la emergencia del virus causante del SARS, se han intensificado los estudios del resto de los componentes de la replicasa, que constituyen potenciales dianas

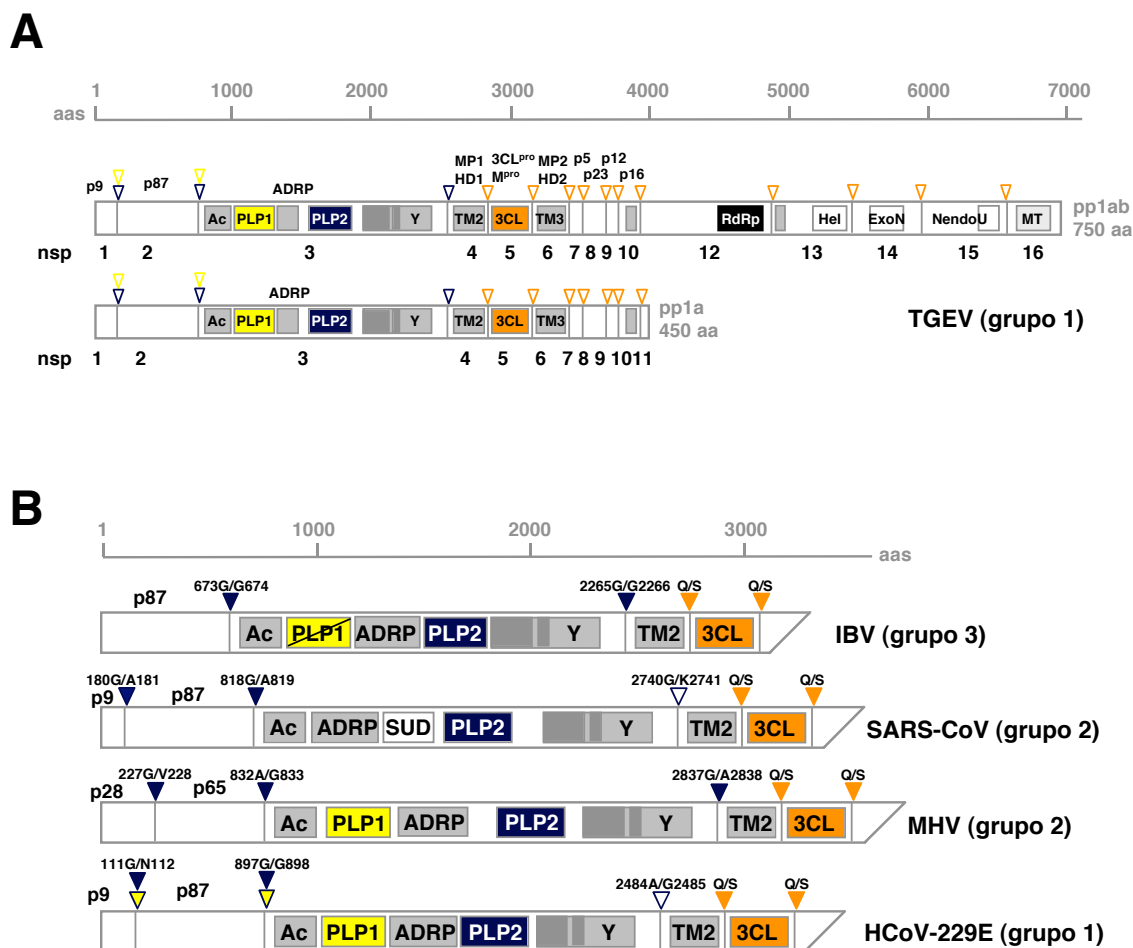


Figura 7. Organización de la replicasa de coronavirus. (A) Organización y procesamiento proteolítico de la replicasa del TGEV (por homología con otros CoV). Los sitios de procesamiento por las proteasas PLP1, PLP2 y 3CL se indican con triángulos invertidos, amarillos, azules o naranjas, respectivamente. Las nsps están numeradas del 1 al 11 en el caso de la poliproteína pp1a y del 1 al 10 y del 12 al 16 en la poliproteína pp1ab. En la parte superior o dentro de los recuadros se indican nombres alternativos para designar los productos finales de procesamiento. P9 y p87, proteínas N-terminales de 9 y 87 kDa, respectivamente; Ac, dominio ácido; PLP1, PLP2, proteasas accesorias similares a papaína 1 y 2; ADPR, actividad adenosina difosfato ribosa 1'' fosfatasa; Y, dominio Y, que contiene una región transmembrana y un dominio potencial de unión a metales; TM1 a TM3, potenciales dominios transmembrana; 3CL, proteasa principal; RdRp, RNA polimerasa dependiente de RNA; Hel, helicasa; ExoN, exoribonucleasa 3'-5'; NendoU, endoribonucleasa; MT, ribosa 2'-O-metiltransferasa. (B). Procesamiento proteolítico N-terminal para los tres grupos de coronavirus. El virus IBV contiene una forma inactiva de la proteasa PLP1 (tachada por una línea diagonal) y el virus SARS-CoV carece de esta enzima. Los sitios de procesamiento se indican con triángulos invertidos, con el código de color de cada proteasa o sin colorear (sitios de procesamiento no demostrados). SUD, dominio único en el SARS-CoV.

terapéuticas para controlar la infección por coronavirus. Recientemente se ha descrito una colección de enzimas de procesamiento de RNA que no están presentes en la mayoría de los virus RNA (Snijder y col., 2003) y que incluyen una endoribonucleasa (nsp15) (Renzi y col., 2006; Ricagno y col., 2006), una 3'-5' exoribonucleasa (nsp14) (Minskaia y col., 2006), una ribosa 2'-O-metiltransferasa (nsp16) (Ginalski y col., 2006; von Grotthuss y col., 2003); y una ADP ribosa 1'' fosfatasa (nsp3) (Putics y col., 2006). La función de estas enzimas durante la síntesis de RNA en coronavirus es desconocida, aunque se postula que juegan un papel esencial en el proceso de transcripción discontinua y en la replicación del genoma. La presencia de la actividad 3'-5' exoribonucleasa diferencia a los nidovirales con genomas próximos a 30 kb, de los arterivirus, que contienen genomas en torno a 14 kb. Esta exoribonucleasa se ha postulado que podría formar parte de un sistema de corrección de errores de la RdRp junto con las otras enzimas de procesamiento del RNA, que sería necesario para mantener la estabilidad del genoma por encima de un tamaño crítico (Gorbalenya y col., 2006).

Las proteínas nsp3, nsp4 y nsp6 contienen dominios hidrofóbicos, posiblemente responsables de la asociación del complejo de replicación en las membranas intracelulares. En el extremo C-terminal de la poliproteína pp1a se encuentra un grupo de proteínas pequeñas nsp7-nsp10 de las que no se tienen datos funcionales, con la excepción de la nsp8, descrita recientemente como una nueva polimerasa dependiente de RNA. Esta enzima presenta una actividad similar a las “primasas” celulares y podría sintetizar RNAs cortos que serían utilizados por la RdRp como iniciadores (Imbert y col., 2006).

7. PROTEÍNAS CELULARES IMPLICADAS EN LA REPLICACIÓN DE CORONAVIRUS

Los virus RNA, debido a su capacidad genética limitada, dependen de la célula hospedadora en muchos pasos durante la síntesis de sus macromoléculas, y a menudo secuestran y redirigen factores celulares para su propia traducción, replicación y transcripción.

La identificación de proteínas celulares implicadas en la replicación de virus RNA se ha llevado a cabo siguiendo diferentes aproximaciones: (i) mediante estudios de permisividad de determinados tipos celulares para la replicación viral; (ii) analizando

los factores que se unen específicamente o copurifican con componentes de la replicasa viral; (iii) estudiando las proteínas que interaccionan con el genoma, preferentemente con los extremos 5' y 3'; (iv) realizando análisis mutacionales en sistemas genéticos modelo como *Sacharomyces cerevisiae* y, más recientemente, (v) mediante el uso de técnicas de silenciamiento génico. En algunos casos los factores identificados se han correlacionado funcionalmente con la replicación viral, aunque en muchos otros casos no se ha establecido una relación causal directa y necesitan de una verificación posterior. Por lo general, las proteínas que se han implicado en la síntesis de RNA viral son de naturaleza muy diversa e incluyen proteínas de procesamiento del RNA, factores de transcripción, factores de traducción, proteínas del citoesqueleto, chaperonas y diversas proteínas de membrana (Ahlquist y col., 2003; Boguszewska-Chachulska y Haenni, 2005; Lai, 1998).

Un ejemplo clásico de proteínas celulares asociadas a complejos de replicación de virus RNA es el del fago Q β , cuya replicasa está formada por la RdRp viral, los factores de elongación de la traducción Tu (EF-Tu) y Ts (EF-Ts) y la proteína ribosomal S1 (Blumenthal y Carmichael, 1979).

Uno de los sistemas de replicación mejor caracterizados para un virus RNA de cadena positiva es el caso de poliovirus, donde se han identificado proteínas celulares de unión a RNA implicadas en su traducción y replicación, como son las proteínas de unión a secuencias de poliA, o a secuencias ricas en citidinas (PABP y PCBP, respectivamente), o la nucleolina (Gamarnik y Andino, 1997; Herold y Andino, 2001; Parsley y col., 1997; Waggoner y Sarnow, 1998; Walter y col., 2002).

En coronavirus, se ha propuesto la participación de varios factores celulares en la síntesis del RNA viral. El descubrimiento de estas proteínas se ha realizado en base a su capacidad de interaccionar con diversos motivos de RNA de importancia funcional de los virus BCoV y MHV. Una de las proteínas más estudiadas es la proteína heterogénea nuclear A1 (hnRNP A1), identificada inicialmente por su interacción con la TRS de cadena negativa de MHV (Furuya y Lai, 1993; Li y col., 1997; Zhang y Lai, 1995). La hnRNP A1, interacciona con alta afinidad con la proteína N viral (Wang y Zhang, 1999), con la 3'-UTR de cadena positiva (Huang y Lai, 2001) y es capaz de dimerizar. Estas propiedades junto con su función celular como modulador de *splicing*, hacen de la hnRNP A1 un buen candidato para mediar la aproximación de secuencias distales en el genoma durante la síntesis de RNA de coronavirus. Existen datos experimentales que

sugieren un papel importante de la hnRNP A1 en la replicación de coronavirus. La sobreexpresión de la hnRNP A1 favorece la síntesis de RNA en MHV, mientras que la expresión de una forma truncada de la proteína tiene un efecto de dominante negativo sobre la replicación (Shi y col., 2000). Adicionalmente, la proteína se relocaliza en el citoplasma de células infectadas por MHV, donde ocurre la síntesis de RNA (Li y col., 1997) y la mutación de los sitios de unión de la hnRNP A1 reduce la replicación y transcripción de RNAs defectivos de MHV (Huang y Lai, 2001). Por otro lado, el papel de la hnRNP A1 se ha cuestionado en base a la replicación eficiente de MHV en una línea celular que no expresa niveles detectables de la proteína (Ben-David y col., 1992; Shen y Masters, 2001), aunque existen varias isoformas de la hnRNP A1 que probablemente puedan reemplazar su función en esta línea celular (Shi y col., 2003).

Estudios clásicos de retardo en gel han permitido localizar sitios de unión de factores celulares con RNAs del extremo 3' de los coronavirus MHV y BCoV, donde residen señales de replicación conservadas, aunque muchas de estas proteínas no han podido ser identificadas. Distintas proteínas, con masas moleculares de 99, 95, 73, 40-50 y 30 kDa se unen a la secuencia 3'UTR de BCoV, de las cuales la proteína de 73 kDa se ha identificado como la proteína de unión a tramos de poliadenina (PABP). Esta interacción se conservaba en el coronavirus MHV y correlaciona funcionalmente con el requerimiento de la cola de poliA para la replicación de RNAs defectivos derivados de estos virus (Spagnolo y Hogue, 2000). Por otro lado, los últimos 42 nt del extremo 3' de MHV interaccionan con proteínas de 90, 70, 58 y 40 kDa, de las que se identificó la proteína de 90 kDa como la aconitasa mitocondrial. Esta proteína se observó que colocalizaba con la proteína N en los sitios de replicación de MHV y se encontró una correlación entre sus niveles de expresión y el crecimiento del virus (Nanda y Leibowitz, 2001). Actualmente, los avances en proteómica podrían permitir la identificación sistemática de las proteínas celulares que interaccionan con las señales de replicación de coronavirus.

Otra proteína de la familia de las hnRNPs identificada por su interacción con secuencias de coronavirus es la proteína de unión a trectos de polipirimidinas (PTB o hnRNP I), que se une de forma específica con la cadena positiva de la secuencia líder de MHV (Li y col., 1999) y con la cadena complementaria a la 3'UTR (c3'UTR) (Huang y Lai, 1999). La mutación de los sitios de interacción de la PTB afecta negativamente a la síntesis de RNA de MHV. Interesantemente, se observó que las proteínas hnRNP A1 y

PTB se unían a secuencias complementarias en ambos extremos del genoma de MHV y que ambas proteínas recombinantes eran capaces de formar un complejo *in vitro* con fragmentos de los extremos 5' y 3' del genoma (Huang y Lai, 2001), por lo que se han implicado en una potencial interacción de los extremos. Sin embargo, esta interacción, propuesta para MHV, no ha sido demostrada en presencia de proteínas endógenas, y se desconoce su conservación entre otros coronavirus, así como su relevancia funcional.

La circularización del genoma es una característica común para un número creciente de virus RNA, ya sea mediada por interacciones RNA-RNA a larga distancia como se ha propuesto en flavivirus (You y col., 2001) o mediada por interacciones RNA-proteína y proteína-proteína, como ocurre en rotavirus (Piron y col., 1998) y picornavirus (Herold y Andino, 2001). Se ha propuesto que esta interacción puede tener un papel en la traducción, transcripción o replicación de estos virus. Este mecanismo, mimetiza el de la traducción de los RNAs mensajeros celulares, que tiene lugar según el llamado modelo de bucle cerrado o *closed-loop model*, conformación que facilita la regulación positiva de la traducción (Bergamini y col., 2000; Craig y col., 1998; Michel y col., 2001; Otero y col., 1999; Tarun y Sachs, 1996).

IV. OBJETIVOS

Los coronavirus son importantes patógenos animales y humanos, cuyo mecanismo de replicación es prácticamente desconocido. Estos virus poseen los genomas de mayor longitud entre los virus RNA y una replicasa de elevada complejidad, que aumentan el interés del estudio de su replicación.

El objetivo general de esta tesis es el estudio de ciertos requerimientos de la replicación de coronavirus, utilizando como modelo el virus TGEV. Los objetivos específicos abordados en este trabajo fueron los siguientes:

1. Determinar los requerimientos de una mutación puntual en el gen de la replicasa, localizada en el sitio de procesamiento nsp1/nsp2, sobre la replicación del virus TGEV y la amplificación de minigenomas derivados del mismo.

2. Estudiar los requerimientos para la replicación autónoma del RNA del TGEV, mediante la delección secuencial de genes estructurales y accesorios del extremo 3' del genoma. Aplicación al desarrollo de vectores virales eficientes en replicación y deficientes en propagación basados en un genoma único.

3. Estudio de la posible interacción entre los extremos 5' y 3' del genoma viral durante la replicación de coronavirus. Analizar la interacción de proteínas virales y celulares con las señales de replicación localizadas en los extremos 5' y 3' del genoma.

V. MATERIALES Y MÉTODOS

1. CÉLULAS EUCARIÓTICAS

1.1. Líneas celulares

La línea celular ST (swine testis), derivada de testículo embrionario de cerdo, fue inicialmente desarrollada por McClurkin y Norman (1966) y cedida al laboratorio por L. Kemeney (National Animal Disease Center, Ames, Iowa, EEUU).

La línea celular BHK-pAPN (Delmas y col., 1992), derivada de la línea epitelial de riñón de hámster BHK-21 (ATCC CCL10) y transformada establemente con el receptor del TGEV, la aminopeptidasa N porcina, fue cedida al laboratorio por el Dr. H. Laude (Unité de Virologie Immunologie Moléculaires, Jouy-en-Josas, Francia).

La línea celular 293T (HEK 293T/17) fue adquirida por el laboratorio a la *American Type Culture Collection* (ATCC CRL-11268).

La línea celular HeLa (ATCC CCL-2), procedente de un carcinoma cervical humano, fue cedida al laboratorio por el Dr. J. Ortín (Centro Nacional de Biotecnología, Madrid, España).

1.2. Cultivo de células eucarióticas

Las células ST, HeLa, BHK-pAPN y 293T se cultivaron a 37°C en una atmósfera de CO₂ al 7% y una humedad del 98% en medio DMEM suplementado con aminoácidos no esenciales al 1% (Sigma), suero de ternera fetal (STF) (Bio-Whittaker Europe, Verviers, Bélgica) al 10%, 2 mM L-glutamina (Sigma) y 50 mg/ml de gentamicina. En el caso de las células BHK-pAPN, establemente transformadas con el receptor del TGEV, cada dos pases se crecieron en presencia del antibiótico de selección geneticina (G418), que se añadió al medio de cultivo a una concentración de 1.5 µg/µl. En todos los casos, la concentración de STF en el medio se redujo al 2% durante las infecciones. Las líneas celulares se almacenaron mediante congelación en nitrógeno líquido a una densidad de 1-2 x 10⁶ células/ml en una mezcla 9:1 de STF:DMSO.

2. VIRUS

2.1. Aislados virales

El aislado de virus TGEV PUR46-MAD (código de acceso AJ271965) (Sánchez y col., 1990) corresponde al clon PUR46-CC120-MAD aislado en el laboratorio, que

deriva del aislado Purdue del TGEV, obtenido en el año 1946 en Indiana (EEUU) por Doyle y Hutchings.

Los virus recombinantes rescatados a partir de la transfección del clon infectivo del TGEV presentan el fondo genético del aislado del TGEV PUR46-MAD con la excepción del gen S, derivado del aislado TGEV PUR46-C11 (Almazán y col., 2000; Sánchez y col., 1999).

2.2. Crecimiento y titulación del virus TGEV

El virus TGEV PUR46-MAD se creció y tituló en células ST según lo descrito previamente (Correa y col., 1988). Para la generación de lotes o *stocks* de virus, las células ST se crecieron en monocapas hasta el 100% de confluencia y se infectaron 24 horas después a una multiplicidad de infección (mdi) de 0.1-1 unidades formadoras de placa (UFP) por célula. El sobrenadante de esta infección se recogió entre las 16-20 horas después de la infección (hdi) y se almacenó a -70°C hasta su utilización.

3. BACTERIAS

3.1. Cepas y cultivos de bacterias

La cepa DH10B de *E. coli* (GibcoBRL), cuyo genotipo es F *mcrA* Δ (*mrr-hsdRMS-mcrBC*) Δ 80*dlacZ* Δ M15 Δ *lacX74* *deoR* *recA1* *endA1* *araD139* Δ (*ara, leu*)7697 *galU* *galK* λ -*rpsL* *nupG*, se utilizó de rutina para el clonaje de DNA.

Los cultivos bacterianos se realizaron en medio líquido LB (Sambrook y Russell, 2001) a partir del cual se preparó medio sólido añadiendo agar a una concentración de 15 g/l, que se utilizó para el aislamiento de colonias bacterianas. El medio fue suplementado con los antibióticos adecuados (150 μ g/ml ampicilina; 12.5 μ g/ml cloranfenicol) cuando fue necesario. Los cultivos en medio líquido se crecieron a 30°C o a 37°C, según se indique.

3.2. Preparación de bacterias competentes

Las bacterias DH10B competentes para electroporar se prepararon a partir de una colonia única crecida en medio sólido, que se amplificó posteriormente en un litro de medio SOB (20 g/l triptona, 5 g/l extracto de levadura, 0.5 g/l NaCl, 0.18 g/l KCl) hasta alcanzar una DO₅₅₀ de 0.7. A continuación el cultivo se enfrió a 4°C durante 20 min, las

bacterias se sedimentaron por centrifugación a 4000g durante 15 min y se lavaron tres veces con una solución previamente enfriada a 4°C de glicerol al 10%, utilizando un volumen del cultivo inicial para el primer lavado y reduciendo el volumen a la mitad en los lavados sucesivos (1 l, 0.5 l y 0.25 l, respectivamente). En cada lavado las bacterias se resuspendieron a 4°C y se centrifugaron a 4000g durante 15 min. El sedimento final se resuspendió en 3 ml de glicerol frío al 10%, se repartió en alícuotas y se congeló a -70°C hasta su utilización.

3.3. Transformación de bacterias

Las bacterias DH10B se transformaron por electroporación con 20-40 ng de DNA diluido en 50 µl de agua. El DNA diluido se calentó a 65°C durante 10 min, se enfrió a 4°C y se mezcló con 50 µl de las bacterias competentes. Esta mezcla se incubó durante 1 minuto a 4°C y se transfirió a una cubeta de electroporación de 0.2 cm (Bio-Rad), donde se aplicó un pulso de 25 µF, 2.5 kV y 200 Ω utilizando un electroporador *GenePulser* (Bio-Rad). Las bacterias se resuspendieron inmediatamente en 1 ml de medio SOC (Hanahan, 1985), se incubaron a 37°C con agitación durante 45 min y se sembraron por extensión en superficie sobre placas de medio LB-agar con el antibiótico apropiado.

4. MANIPULACIÓN Y ANÁLISIS DEL DNA

4.1. Plásmidos de clonaje

Los plásmidos pSL1190 (Pharmacia Biotech) y pcDNA3 (Invitrogen), de alto número de copia, se utilizaron para el clonaje y modificación de cDNAs virales de hasta 9.7 kb.

El plásmido pBeloBAC11 (Wang y col., 1997), de 7.5 kb, contiene el origen de replicación del factor F de *E. coli* (*oriS*), el gen de resistencia a cloranfenicol (CM^r) y los genes necesarios para mantener una copia única del plásmido por célula (*parA*, *parB*, *parC* y *repE*). Este vector fue utilizado para el clonaje y modificación de cDNAs virales de gran tamaño o que contuvieran secuencias tóxicas para su crecimiento en bacterias según lo descrito anteriormente (Almazán y col., 2000; González y col., 2002).

4.2. Manipulación y purificación de plásmidos

Para la purificación de plásmidos a partir de cultivos bacterianos a pequeña escala (Minipreps) se utilizó el reactivo *Plasmid Mini Kit* (QIAGEN) según el protocolo del fabricante, con las modificaciones apropiadas en función del tamaño y número de copia del plásmido. Para la purificación de plásmidos de alto número de copia a media y gran escala (Midi- y Maxipreps) se utilizaron los reactivos *Plasmid Midi kit* y *Plasmid Maxi kit* (QIAGEN), respectivamente. Los BACs se purificaron utilizando el reactivo *Large-Construct kit* (QIAGEN), a partir de cultivos bacterianos que se crecieron a 30°C durante 18 h aproximadamente, partiendo de un preinóculo de 10 ml de cultivo que se añadió a 500 ml de medio LB que contenía 12.5 µg/ml de cloranfenicol.

La purificación de DNAs a partir de reacciones enzimáticas se realizó mediante extracción con fenol/cloroformo/alcohol isoamílico (25:24:1, pH 8) y precipitación con 0.1 volúmenes de 3 M NaOAc pH 4.5 y 2.5 volúmenes de 100% EtOH. Después de un lavado con EtOH al 70%, el precipitado de DNA se dejó secar durante 5 min a temperatura ambiente y se resuspendió en TE (10 mM Tris-HCl pH 7.5, 1 mM EDTA).

Para la purificación de fragmentos de DNA extraídos de geles de agarosa y de productos amplificados mediante reacción en cadena de la polimerasa (PCR) se utilizó el reactivo *GeneClean Turbo kit* (Q-BIOgene) o el reactivo Qiaex II (QIAGEN) en el caso de fragmentos mayores de 10 kb, siguiendo el protocolo indicado por el fabricante en cada caso.

4.3. Enzimas de restricción y modificación de DNA

Las enzimas de restricción y modificación de DNA (DNA ligasa del fago T4, fosfatasa alcalina de gamba y el fragmento Klenow de la DNA polimerasa I) se adquirieron a Roche o a New England Biolabs. Los tratamientos enzimáticos de restricción, desfosforilación y ligación de DNA se hicieron siguiendo protocolos estándar (Sambrook y Russell, 2001). Todas las construcciones generadas se analizaron por secuenciación utilizando un secuenciador automático 373 (Applied Biosystems).

4.4. Amplificación del DNA mediante PCR

Las reacciones de PCR con fines analíticos se realizaron utilizando 2.5 U de la enzima *AmpliTaQ DNA polymerase* (Applied Biosystems) en un volumen de reacción de 50 µl en presencia de 20 pmoles de los oligonucleótidos correspondientes, una mezcla

de deoxinucleótidos trifosfato a una concentración final de 0.25 mM, 1.25 mM de MgCl₂ y tampón de PCR (10 mM Tris-HCl pH 8.3, 50 mM KCl). En función de la temperatura de fusión de los oligonucleótidos (T_m) y la longitud del fragmento a amplificar, la reacción de PCR se programó según los ciclos: (a) 94°C, 5 min; (b) 30-35 ciclos de 94°C, 1min; (T_m-10°C), 1 min; 72°C, 1 min/kb; (c) 72°C, 10 min.

Para la amplificación de fragmentos de DNA con fines preparativos se utilizó la enzima *Platinum Pfx DNA polimerase* (Invitrogen), que posee actividad exonucleasa correctora 3'-5', según las indicaciones del proveedor.

4.5. Electroforesis de DNA en geles de agarosa

La separación de DNA para estudios analíticos y purificación de bandas de DNA se realizó en geles de agarosa *D-1 Media EEO* (Pronadisa) al 0.65-1.2% disuelta en TAE (40 mM Tris-acetato, 1 mM EDTA) con 1 µg/ml de bromuro de etidio. Para la resolución y purificación de bandas de pequeño tamaño (< 200 pb) se empleó la agarosa de bajo punto de fusión *NuSieve GTG Agarose* (Cambrex).

5. CONSTRUCCIÓN Y MUTAGÉNESIS DE PLÁSMIDOS

5.1. Mutagénesis de cDNAs del genoma del TGEV

Para la introducción de mutaciones puntuales en la posición 637 del genoma viral se utilizó el plásmido intermedio pBAC-TGEV(*SrfI-NheI*) (Zúñiga y col., 2004) que comprende la secuencia viral entre los nucleótidos 1 al 15062, con la excepción del fragmento tóxico *ClaI-ClaI* que comprende los nucleótidos 4417 a 9615. Las mutaciones puntuales se generaron mediante PCRs solapantes amplificadas con los oligonucleótidos indicados en la tabla III (Fig. 8). El producto final se digirió con las enzimas de restricción *SrfI* y *ApaLI* y se clonó en los mismos sitios de restricción del plásmido pBAC-TGEV (*SrfI-NheI*), generando los plásmidos pBAC-TGEV (*SrfI-NheI*)637A, pBAC-TGEV(*SrfI-NheI*)637T y pBAC-TGEV(*SrfI-NheI*)637C. Para introducir las mutaciones en el cDNA del TGEV, el fragmento *SrfI-ClaI* del plásmido pBAC-TGEV(*SrfI-NheI*) con la mutación correspondiente, se clonó en los mismos sitios del plásmido pBAC-TGEV^{Δ*ClaI*}. Finalmente se reconstruyó el cDNA del TGEV mediante inserción del fragmento *ClaI-ClaI* como ha sido descrito (Almazán y col., 2000).

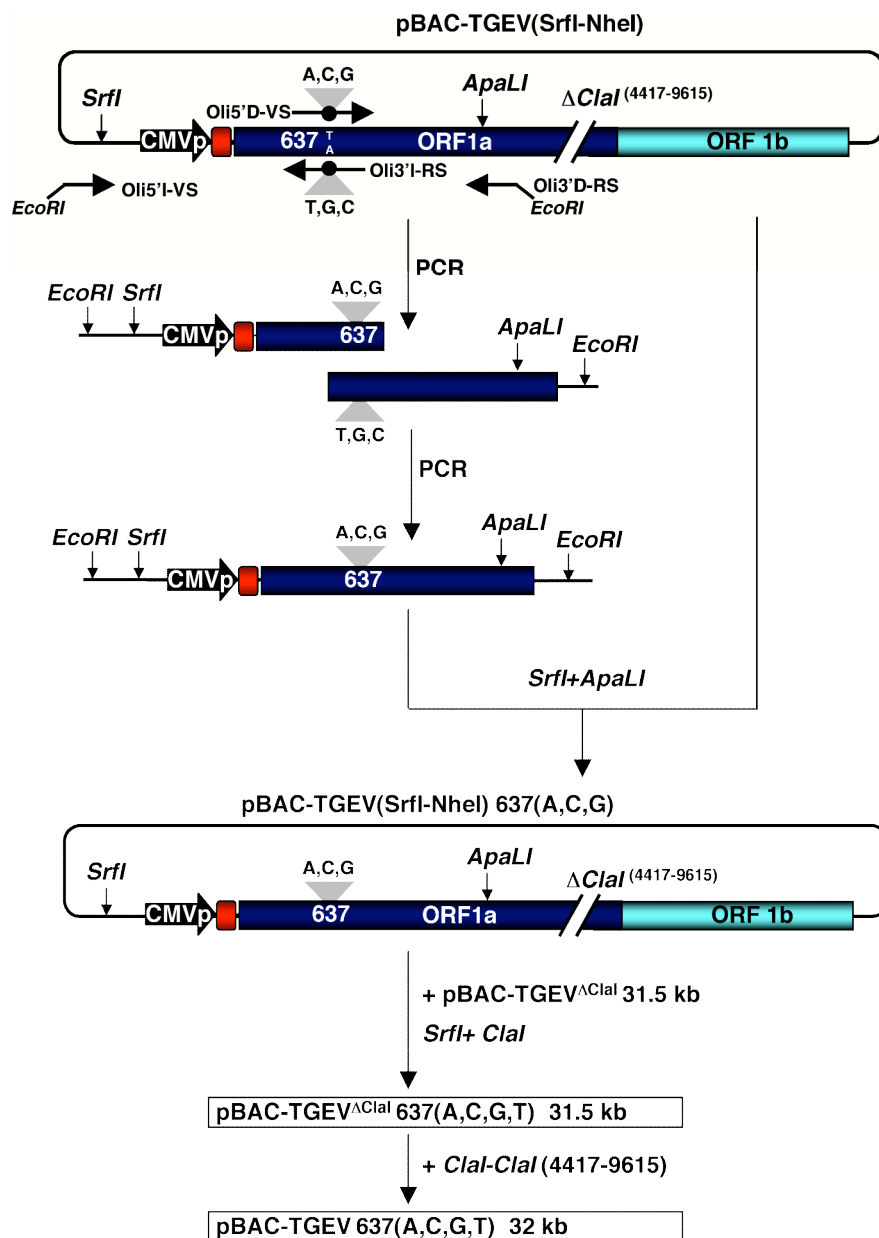


Figura 8. Mutagénesis puntual del genoma del TGEV. Esquema de la estrategia utilizada para la introducción de mutaciones en la posición 637 del genoma viral. ORF1a y ORF1b, gen de la replicasa. La secuencia de los oligonucleótidos se indica en la tabla III. CMVp, promotor de CMV

5.2. Modificación y mutagénesis de minigenomas derivados del TGEV

Para la generación del plásmido pM33L-EGFP, el fragmento de restricción *Sma*I-*Not*I extraído del plásmido pIRES2-EGFP (Clontech), que contiene la secuencia codificante de la proteína verde fluorescente intensificada (EGFP) precedida del IRES del virus de la encefalomiocarditis, se rellenó en los extremos con el fragmento Klenow de la DNA polimerasa I y se insertó en el plásmido pM33L (Izeta, 2000) linearizado con *Nru*I.

TABLA III. OLIGONUCLEÓTIDOS UTILIZADOS PARA LA MUTAGÉNESIS DE LA POSICIÓN 637 DEL GENOMA EN EL PLÁSMIDO pBAC-TGEV(*SrfI*-*NheI*)

Oligonucleótido ^(a)	Secuencia 5' → 3' ^(b)	Posición ^(c)
Oli5'I-VS ^(d)	AATTCGATGATAAGCTGTCAAC	-
Oli3'I-RS	GGCACCTCTG(G/C/T)CAGTGCGAGCAATTTTAAG	647-617
Oli5'D-VS	GCTCGCACTG(A/C/G)CAGAAGGTGCCATATATGTTG	527-657
Oli3'D-RS	TCCTCTACTACTTTCCAAGCGTC	1509-1531

^(a) VS (*virus sense*), sentido viral o polaridad positiva, RS (*reverse sense*), sentido antiviral o polaridad negativa

^(b) Las bases entre paréntesis indican el nucleótido mutado

^(c) Posición en el genoma del TGEV. En los oligonucleótidos de sentido inverso se indica la posición de la secuencia complementaria

^(d) Oligonucleótido complementario al vector en posición anterior al sitio de restricción *SrfI*

Para la generación del plásmido pDIC-EGFP, el mismo fragmento *SmaI*-*NotI* del vector pIRES2-EGFP se introdujo en el plásmido pDI-C (Izeta y col., 1999), que contiene el cDNA del genoma defectivo DI-C, en el lugar originado después de la delección del fragmento *StuI*-*StuI* de 711 pb.

El minigenoma M33L (Fig. 10) presentaba una T en la posición 637 que se sustituyó por una G, que es la secuencia consenso viral, mediante intercambio de un fragmento de restricción. Para ello, el fragmento *NaeI*-*SphI* de 1047 pb se extrajo del plásmido pBAC-TGEV-5'END, que comprende los primeros 1496 nt del genoma viral, y se clonó en el plásmido pM33L digerido con *NaeI* parcialmente y con *SphI*.

Los minigenomas M33L-637A y M33L-637C se generaron a partir de los plásmidos pBAC-TGEV(*SrfI*-*NheI*)637A y pBAC-TGEV(*SrfI*-*NheI*)637C por intercambio del fragmento de restricción *NaeI*-*EcoRV* de 646 pb en el plásmido pM33L digerido con *NaeI* parcialmente y con *EcoRV*.

A partir de las construcciones del minigenoma M33L con las mutaciones en la posición 637 (A, C o G) se generaron los correspondientes mutantes del genoma defectivo DI-C (Fig. 10) por intercambio del fragmento de restricción *NarI*-*EcoRV* de 1229 pb extraído del plásmido pM33L con la mutación adecuada, en el vector derivado del plásmido pDI-C digerido con estas enzimas.

Para la construcción del minigenoma no replicativo M33L-Δ171 (Fig. 14), el fragmento *PmlI*-*Asp700* de 171 nt, que elimina la secuencia comprendida entre las

posiciones 81-252 de la 3'UTR (273 pb), se delecionó del plásmido pM33L-637G.

El minigenoma M39-GUS, utilizado para ensayar la función en *trans* del replicón TGEV-REP, ha sido descrito anteriormente (Alonso y col., 2002). En esta construcción, el minigenoma M39, derivado del RNA defectivo DI-C, está precedido del promotor de CMV y contiene el gen trazador de la β -glucuronidasa bajo la TRS del gen N. Este minigenoma contiene en su extremo 3' la secuencia de la ribozima del virus HDV y las secuencias de terminación y poliadenilación del gen de la hormona BHG, para que el extremo 3' del minigenoma sea idéntico al del RNA viral (Fig. 20A).

5.3. Clonaje del dominio proteasa PLP1 y del sustrato N-terminal de la replicasa

Para la construcción del pásmido pTGEV-PLP1, que codifica el dominio correspondiente a la proteasa PLP1, la región del genoma viral comprendida entre los nucleótidos 2889 a 4259, ambos inclusive, que codifican los aminoácidos Glu859 a Ser1315 de la replicasa, fue amplificada por PCR a partir del plásmido pBAC-TGEV(*SrfI-NheI*) utilizando los oligonucleótidos *KpnI*-M-E859-VS y *BamHI*-STOP-S1315-RS (Tabla IV), los cuales introducen un sitio *KpnI* y un codon de iniciación de la

TABLA IV. OLIGONUCLEÓTIDOS UTILIZADOS PARA EL CLONAJE DE SECUENCIAS DE LA REPLICASA

Oligonucleótido ^(a)	Secuencia 5' → 3' ^(b)	Posición ^(c)
<i>KpnI</i> -M-E859-VS	GCGC <u>GGTACCAT</u> GGAACATTTTATCCTTATGG	2889-2908
<i>BamHI</i> -STOP-S1315-RS	CGCGGGATCCTTAACTTTGAATTTCTCTTTTTC	4239-4259
<i>KpnI</i> -ORF1a-M1-VS	GCGC <u>GGTACCAT</u> GAGTTCCAAACAATTCAAG	315-335
<i>BamHI</i> -STOP-ORF1a-E610-RS	CGCGGGATCCTTATTCAAATGATGAACCAAGTTTGT	2122-2144

^(a) VS (*virus sense*), sentido viral o polaridad positiva, RS (*reverse sense*), sentido antiviral o polaridad negativa

^(b) Las bases subrayadas indican dianas de restricción. Los codones de iniciación (ATG) y de terminación (TAA) de la traducción se indican en negrita. Las secuencias en cursiva se añadieron para facilitar la restricción del fragmento amplificado

^(c) Posición en el genoma del TGEV. En los oligonucleótidos de sentido inverso se indica la posición de la secuencia complementaria

traducción en el extremo 5' y un codon de terminación de la traducción seguido por un sitio *BamHI* en el extremo 3'. Después de la digestión de este fragmento de PCR con las

enzimas *KpnI* y *BamHI*, se clonó en el plásmido pcDNA3 digerido con las mismas enzimas.

Para la generación de los plásmidos pTGEV-ORF1a(637G) y pTGEV-ORF1a(637A), la secuencia correspondiente a los 610 aminoácidos del extremo N-terminal de la replicasa se amplificó por PCR utilizando los oligonucleótidos *KpnI*-ORF1a-M1-VS y *BamHI*-STOP-ORF1a-E610-RS (Tabla IV). Los fragmentos de PCR se digirieron con las enzimas *KpnI* y *BamHI* y se clonaron en el plásmido pcDNA3 digerido con las mismas enzimas.

5.4. Generación de un plásmido no replicativo derivado del clon infectivo del TGEV para optimizar la eficiencia de transfección de DNAs de gran tamaño

El plásmido pBAC-TGEV-NR-EGFP (Fig. 21) se generó con el objetivo de obtener un plásmido no replicativo de tamaño similar al clon infectivo del TGEV para evaluar la eficiencia de transfección. Para ello, el gen de la proteína EGFP, clonado bajo la secuencia IRES del virus de la encefalomiocarditis, se extrajo del plásmido pIRES2-EGFP (Clontech) mediante digestión con las enzimas *NheI* y *NotI*. Este fragmento *NheI*-*NotI* se romizó por tratamiento con el fragmento Klenow de la DNA polimerasa I y se clonó en el sitio *ApaI* del plásmido pBAC-TGEV(*SrfI*-*NheI*)637A, generando la construcción pBAC-TGEV(*SrfI*-*NheI*)IRES-EGFP. A partir de este plásmido se obtuvo el fragmento *SrfI*-*ClaI* que se clonó en el plásmido pBAC-TGEV^{Δ*ClaI*} digerido con las enzimas *SfiI* y *ClaI*. Finalmente el fragmento *ClaI*-*ClaI* se introdujo en la orientación inversa para obtener el plásmido no replicativo pBAC-TGEV-NR-EGFP.

5.5 Construcción de replicones derivados del genoma del TGEV

Para la construcción del replicón mínimo teórico TGEV-REP (Fig. 20), la secuencia del TGEV comprendida entre los nucleótidos 1 a 20372, correspondiente a la región 5' UTR y al gen de la replicasa, y la secuencia entre los nucleótidos 28087 a 28585 correspondiente al extremo 3' del genoma, se clonaron en un BAC según lo descrito anteriormente (Almazán y col., 2000). El cDNA resultante está flanqueado en su extremo 5' por el promotor inmediatamente temprano de CMV y en su extremo 3' por la ribozima del virus HDV y por las secuencias de poliadenilación y terminación del gen de la hormona BGH.

Con el fin de analizar qué genes además de la replicasa eran necesarios para la generación de un replicón funcional del TGEV se construyeron tres replicones (REP 1, REP 2 y REP 3) mediante la deleción sucesiva de los genes del extremo 3' del genoma, a partir de un clon infectivo modificado del TGEV (pBAC-TGEV-RS) en el cual los genes del virus se separaron mediante la duplicación de secuencias solapantes y se insertaron sitios únicos de restricción entre cada par de genes (Fig. 9) (Ortego y col., 2003).

TABLA V. OLIGONUCLEÓTIDOS UTILIZADOS PARA LA GENERACIÓN DE REPLICONES DEL TGEV

Oligonucleótido ^(a)	Secuencia ^(b)
Rep1-VS	5'-CGCGT <u>TAA</u> TAAGAAGGGTAAGTTGCTCATTAGAAATAATG GTAAGTTATTGAATTTTGGTAACCACTTCGTTAACAC-3'
Rep1-RS	5'-GGAGTGGCCGGCC <u>TATTTAAATCTGACGCGTCTCACCTA</u> GGTTCAAGGTGTGTTAACGAAGTGGTTACCAAAATTCAA-3'
<i>EcoRV-PmeI</i> -TRS3a-VS	5'-GCGGATATCTGTTTAAACGTGTGGCTACTAATAGGCTTAGTAG-3'
<i>EcoRV-SfoI</i> -EGFP-RS	5'-GGGGATATCGGCGCCAATATTTAAATAAAATCACCATTGAG-3'

^(a) VS (*virus sense*), sentido viral o polaridad positiva, RS (*reverse sense*), sentido antiviral o polaridad negativa

^(b) Las bases subrayadas indican dianas de restricción. Las mutaciones e inserciones de secuencia no viral se indican en negrita. Las secuencias en cursiva se añadieron para facilitar la restricción del fragmento amplificado

El plásmido pBAC-TGEV-REP1 contiene el cDNA del replicón REP 1 del TGEV, que incluye el gen de la replicasa, los extremos 5' y 3' del genoma y los genes M, N y 7. Esta construcción se generó a partir del plásmido pBAC-TGEV-RS^{ΔC_{la}I}, que se digirió con las enzimas *PacI* y *FseI*, eliminando un fragmento de 5819 pb que contiene los genes S, 3a, 3b y E (Fig. 9). En su lugar se clonó un fragmento sintético de 127 pb generado por PCR a partir de los oligonucleótidos solapantes Rep1-VS y Rep1-RS (Tabla V), que contiene los sitios de restricción *PacI*, *AvrII*, *MluI*, *SwaI* y *FseI* (Fig. 9). La CS y el ATG del gen de la S, que se encuentran al final de la secuencia del gen de la replicasa, se modificaron mediante la introducción de mutaciones silenciosas en el fragmento sintético, con el fin de evitar la posible interferencia en la expresión de genes heterólogos que se quieran clonar utilizando el sitio múltiple de clonaje generado.

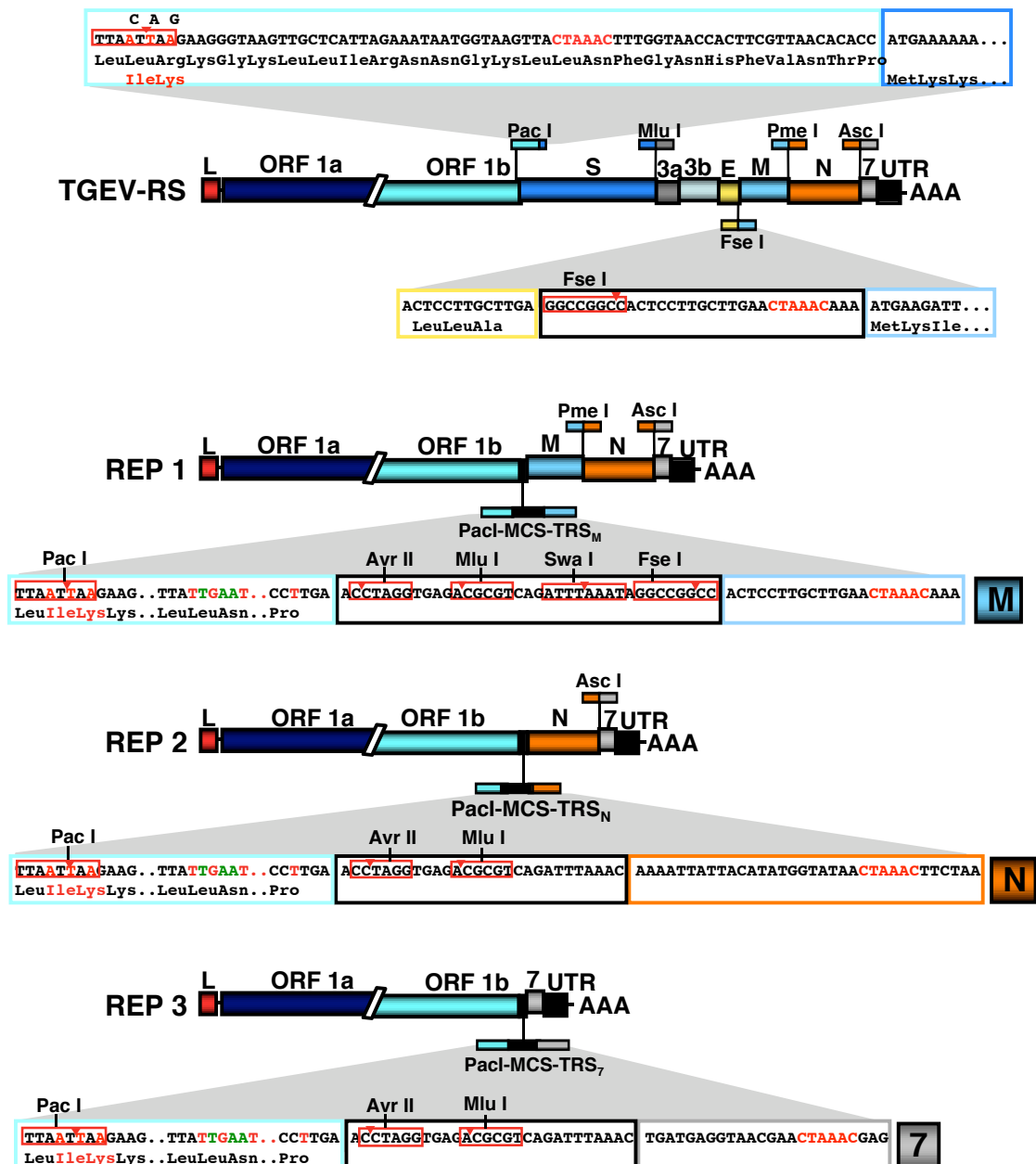


Figura 9. Estructura genética de los replicones derivados del TGEV. Esquema de la estructura del clon infeccioso del TGEV que contiene sitios únicos de restricción entre cada par de genes consecutivos (TGEV-RS) y los replicones generados a partir de esta construcción mediante delección sucesiva de genes del extremo 3' del genoma (REP 1, REP 2 y REP 3). Se muestran en cada construcción los sitios únicos de restricción relevantes. Para evitar posibles interferencias con la expresión de genes heterólogos se introdujeron mutaciones puntuales (nucleótidos en verde) en la TRS y el ATG del gen S. Las secuencias se representan en cajas con el mismo código de color que los genes del virus. Las letras y números sobre las barras indican los genes virales, L, secuencia líder; UTR, región no traducida; AAA, poliA; MCS, sitio múltiple de clonaje.

Finalmente, el fragmento *ClaI-ClaI* se insertó en el último paso de clonaje, generando el plásmido pBAC-TGEV-REP1. Los plásmidos pBAC-TGEV-REP2 y pBAC-TGEV-REP3 se generaron a partir del pBAC-TGEV-Rep1^{Δ*ClaI*}. El plásmido pBAC-TGEV-Rep2^{Δ*ClaI*} se generó por digestión con *SwaI* y *PmeI* y posterior religación. El plásmido pBAC-TGEV-Rep3^{Δ*ClaI*} se construyó mediante digestión con *SwaI* y *AscI*, romización con el fragmento Klenow de la DNA polimerasa I y posterior religación. Finalmente, el fragmento *ClaI-ClaI* se insertó en el último paso de clonaje dando lugar a las construcciones pBAC-TGEV-REP2 y pBAC-TGEV-REP3.

Para analizar la capacidad de expresión de genes heterólogos a partir de los replicones derivados del TGEV, el gen de la EGFP se clonó bajo el control de la TRS del gen 3a en el replicón REP 2. Para ello, el gen de la EGFP precedido de la TRS del gen 3a se amplificó por PCR a partir de la construcción pBAC-TGEV-Δ3ab-EGFP (un clon infectivo del TGEV modificado que expresa el gen de la EGFP bajo el promotor del gen viral 3a) utilizando los oligonucleótidos *EcoRV-PmeI*-TRS3a-VS y *EcoRV-SfoI*-EGFP-RS (Tabla V). El producto de PCR resultante, TRS3a-EGFP, de 1115 bp fue digerido con *EcoRV* y clonado en el plásmido pBAC-TGEV-REP2^{Δ*ClaI*} digerido con la enzima *MluI* y romizado con Klenow, para dar lugar al plásmido pBAC-TGEV-REP2^{Δ*ClaI*}-TRS3a-EGFP. El plásmido pBAC-TGEV-REP2-TRS3a-EGFP se generó mediante la inserción final del fragmento *ClaI-ClaI* (Fig. 28).

5.6. Mutagénesis del gen N del TGEV en el replicón del virus Sindbis

El gen N del TGEV fue clonado en el plásmido pSINrep21, cedido por el Dr C. M. Rice al laboratorio, dando lugar a la construcción pSINrep21-N. La construcción pSINrep21 contiene un replicón del alfavirus Sindbis, que presenta mutaciones adaptativas que le confieren la capacidad de mantener su expresión persistente en células BHK (Frolov y col., 1999). Este vector viral expresa desde un RNA subgenómico el gen *pac* que codifica la enzima puromicina N-acetiltransferasa (PAC) de *Streptomyces alboniger*. La mayoría de las células son susceptibles a la puromicina, que es un aminoacilnucleósido inhibidor de la síntesis de proteínas en procariotas y eucariotas, lo cual permite la selección de las células que contienen el replicón de Sindbis en presencia de puromicina. El gen N del TGEV se clonó en el sitio *MluI* del plásmido pSINrep21 dando lugar al plásmido pSINrep21-N (Fig. 29A). Partiendo de este plásmido, se realizó la mutagénesis puntual del gen de la proteína N mediante

PCRs solapantes. Para ello, se diseñaron oligonucleótidos para mutar la tercera posición del triplete que codifica cada una de las tres metioninas de la proteína N (ATG por ATC) y oligonucleótidos exteriores que contienen la diana de restricción *Mlu*I en su extremo 5', para la inserción de los fragmentos de PCR en el vector pSINrep21 digerido con esta enzima (Tabla VI). Se generó una primera construcción conteniendo la mutación de la metionina iniciadora del gen N (pSINrepN1). Sobre éste plásmido se introdujo la mutación de la segunda metionina (pSINrepN2), y sobre éste se mutó la tercera metionina (pSINrepN3).

TABLA VI. OLIGONUCLEÓTIDOS UTILIZADOS PARA LA MUTAGÉNESIS DEL GEN N

Oligonucleótido ^(a)	Secuencia 5' → 3' ^(b)	Posición ^(c)
<i>Mlu</i> I(-M1)N-VS	TAGAC <u>CGCT</u> ATCGCCAACCAGGGACAACG	26917-26936
(-M2)N-VS	ACTCGCTATCGCAT <u>CGT</u> GAAGGGCCAACG	27135-27164
(-M2)N-RS	CGTTGGCCCTTCACGATGCGATAGCGAGTTTG	27133-27163
(-M3)N-VS	AAGGATGGTGCCATCAACAAACCAACCACG	27264-27294
(-M3)N-RS	TGGTTGGTTTGTGATGGCACCATCCTTGG	27263-27293
<i>Mlu</i> I-N-RS	GTGAGATCTAC <u>GCGT</u> TAGTTCGTTACC	28053-28065

^(a) VS (*virus sense*), sentido viral o polaridad positiva, RS (*reverse sense*), sentido antiviral o polaridad negativa. El orden de la metionina cuyo codon está mutado en cada caso se indica como M1, M2 o M3

^(b) Las bases subrayadas indican dianas de restricción. Las mutaciones se indican en negrita

^(c) Posición en el genoma del TGEV. En los oligonucleótidos de sentido inverso se indica la posición de la secuencia complementaria

5.7. Construcción de plásmidos para la transcripción *in vitro* de los extremos del genoma del TGEV

El plásmido que contiene el extremo 5' del genoma del TGEV se generó a partir del plásmido pM33L-637G. El minigenoma contenido en este plásmido está precedido por el promotor del fago T7 seguido de tres guaninas desde la posición +1 del inicio de la transcripción para su expresión eficiente y una C como primer nucleótido viral. Para la generación de un RNA con el menor número de nucleótidos adicionales en el extremo 5' del virus, el promotor de T7 se substituyó por el del fago SP6 que permite un buen rendimiento de transcripción con una única G en el sitio +1 de la transcripción. Adicionalmente, el primer nucleótido de la secuencia viral se substituyó por una A según la secuencia viral consenso. Estas modificaciones se introdujeron mediante un

fragmento de PCR. Para ello, un fragmento de 1050 pb del extremo 5' del genoma viral se amplificó utilizando como molde el plásmido pM33L-637G y los oligonucleótidos *Xma*I-SP6pG-5'TGEV-VS y TGEV-1043-RS (Tabla VII). El fragmento resultante se digirió con las enzimas *Xma*I y *Eco*RV para su clonaje en el vector pM33L-637G digerido con estas mismas enzimas y tratado con fosfatasa alcalina, generando la construcción pM33L-SP6-5'TGEV. Este plásmido se linearizó alternativamente con las enzimas *Eco*RV o *Bgl*II y se purificó mediante extracción fenólica y precipitación con AcONa/EtOH. Finalmente los plásmidos linearizados pM33L-SP6-5'TGEV-*Eco*RV y pM33L-SP6-5'TGEV-*Bgl*II se resuspendieron a una concentración de 1 µg/µl en 10 mM Tris-HCl pH 7.5 preparado con agua tratada con DEPC y se utilizaron como moldes de transcripción para la generación de RNAs del extremo 5' del virus de 948 nt y 1656 nt, respectivamente. Para la síntesis de un RNA de 0.5 kb del extremo 5' del virus se amplificó un fragmento de PCR comprendiendo los 504 primeros

TABLA VII. OLIGONUCLEÓTIDOS UTILIZADOS PARA LA GENERACIÓN DE PLÁSMIDOS QUE CONTIENEN LOS EXTREMOS 5' Y 3' DEL GENOMA

Oligonucleótido ^(a)	Secuencia ^(b)	Posición ^(c)
<i>Xma</i> I-SP6pG-5'TGEV-VS	5' <i>CGCG</i> <u>CCCCGGG</u> (ATTTAGGTGACACTATAG)AC	1-25
TGEV-1043-RS	TTTTAAAGTAAAGTGAGTGTAGC-3'	
5'TGEV-504-RS	5'-ACCACACGGGCATCTAAGAGTAGCAGCG-3'	1016-1043
	5'-CACCAATGACGTAGTGATCCTTACG-3'	480-504
<i>Xma</i> I-T7pGGG-ORF7-VS	5'- <i>CGCG</i> <u>CCCCGGG</u> (TAATACGACTCACTATAGGG)A	28087-28112
	TGCTGTATTTATTACAGTTTTTAATC-3'	
<i>Hind</i> IIIpolyA-RS	5' <i>GCCGAAGCTT</i> AAAAAAAAAAAAAAAAAAAAAAAAAA	28570-28580
	AAAGTATCACTATC-3'	

^(a) VS (*virus sense*), sentido viral o polaridad positiva, RS (*reverse sense*), sentido antiviral o polaridad negativa

^(b) Las bases subrayadas indican sitios de restricción. Los promotores de transcripción SP6p/T7p están entre paréntesis, el nucleótido de inicio de la transcripción se muestra en negrita. Las secuencias en cursiva se añadieron para facilitar la restricción del fragmento amplificado

^(c) Posición en el genoma del TGEV. En los oligonucleótidos de sentido inverso se indica la posición de la secuencia complementaria

nucleótidos del genoma utilizando los oligonucleótidos *Xma*I-SP6p-5'TGEV-VS y 5'TGEV(504)RS y el plásmido pM33L-SP6-5'TGEV como molde de PCR. El fragmento de PCR fue purificado de la misma forma que los plásmidos anteriores y se utilizó directamente como molde de transcripción.

Para la generación de RNAs del extremo 3' del genoma del TGEV, se construyó el plásmido pT7-3'TGEV, que contenía el promotor del fago T7 seguido de tres guaninas precediendo a la secuencia derivada del gen 7 contenida en el minigenoma M33L. Para ello se amplificó un fragmento de PCR con los oligonucleótidos *Xma*I-T7pGGG-ORF7-VS y *Hind*IIIpolyA-RS (Tabla VII) a partir del plásmido pM33L-637G, se digirió con las enzimas *Xma*I y *Hind*III y se clonó en el vector pM33L-637G digerido con estas mismas enzimas (Fig. 33). El plásmido se linearizó mediante digestión con la enzima *Hind*III y se trató con la nucleasa *Mung Bean* (New England Biolabs) para eliminar los tres nucleótidos 5'-protuberantes adicionales a la secuencia de poliA en la cadena molde de la transcripción. El molde de transcripción resultante se denominó pT7-3'TGEV-*Hind*III-MBN.

6. MANIPULACIÓN Y ANÁLISIS DEL RNA

6.1. Extracción y purificación del RNA intracelular

El RNA total intracelular se extrajo utilizando el reactivo *Rneasy kit* (QIAGEN) según el protocolo del fabricante. El RNA purificado se cuantificó espectrofotométricamente utilizando un espectrofotómetro NanoDrop ND-1000 (NanoDrop Technologies, EEUU) y se guardó en alícuotas a -70°C hasta su utilización.

6.2. Análisis del RNA mediante RT-PCR

La síntesis de la primera cadena de cDNA o retrotranscripción (RT) se realizó partiendo de 200-700 ng de RNA total intracelular utilizando 10 pmoles de oligonucleótido específico complementario a la secuencia a amplificar y 5 U de la transcriptasa reversa M-MLV del retrovirus Moloney de la leucemia murina (Ambion). En un primer paso el RNA se calentó a 70°C durante 5 min y posteriormente se añadió la mezcla de reacción conteniendo el tampón de síntesis de cDNA, 0.05 mM de una mezcla de deoxinucleótidos trifosfato y 0.5 U/μl del inhibidor de RNAsas *RNasin* (Promega) o *SUPERase-In* (Ambion), indistintamente. Las reacciones se incubaron a 42°C durante 1 h en un volumen final de 20 μl.

Una fracción del cDNA resultante de las reacciones de RT (1/10 o 1/5 del volumen de reacción) se utilizó como molde para su amplificación por PCR. Para estudios analíticos, la enzima *AmpliTaq DNA polymerase* (Applied Biosystems) se utilizó en la reacción de PCR según el protocolo descrito en el apartado 4.4 de esta sección.

Alternativamente, para la amplificación con fines preparativos se utilizó la enzima *Platinum Pfx DNA polymerase* (Invitrogen) según las recomendaciones del fabricante.

6.3. Cuantificación del RNA mediante RT-PCR a tiempo real

Los oligonucleótidos empleados para la cuantificación a tiempo real de RNAs virales mediante RT-PCR (Q-RT-PCR) se diseñaron utilizando el programa *Primer Express v2.0* (Applied Biosystems) (Tabla VIII). En todos los casos se comprobó experimentalmente que no formaran dímeros, y se optimizó su concentración en la reacción de PCR entre 300-900 nM. La reacción de RT se realizó en un volumen final de 20 µl con 5 U de la retrotranscriptasa termoestable *Reflectase RT* (Active Motif) en presencia del tampón de síntesis de cDNA comercial de la enzima, 5 mM DTT, 10 pmoles de oligonucleótido específico de sentido inverso y 0.2 mM de una mezcla de deoxinucleótidos trifosfato. Antes de añadir la mezcla de reacción el RNA (entre 50-700 ng) se incubó a 70°C durante 5 min. La reacción de RT se incubó durante 90 minutos a 50°C y la enzima se inactivó mediante incubación a 85°C durante 5 min. Los cDNAs resultantes se guardaron a -20°C hasta su cuantificación.

Para la reacción de PCR cuantitativa se utilizó el reactivo *SYBR Green PCR master mix* (Applied Biosystems) según las especificaciones del fabricante. En cada caso la cuantificación se hizo relativa a una recta estándar de diluciones de orden diez a partir de una reacción de RT de una muestra positiva control. La detección se realizó utilizando un equipo *ABI PRISM 7000* (Applied Biosystems) en un volumen de 25 µl utilizando los parámetros universales de termociclación: (a) 95°C, 10 min; (b) 95°C, 15 seg; 60°C, 1 min; (40 ciclos). A tiempo final de la reacción de PCR, se añadió un paso de rampa de desnaturalización desde 60°C (T_m de los oligonucleótidos) hasta 95°C, para el análisis de la curva de disociación del producto amplificado. Los datos fueron analizados utilizando el programa *ABI PRISM 7000 SDS v.1.0*. Debido a la ausencia de un mensajero celular adecuado para normalizar la cuantificación en las distintas condiciones experimentales utilizadas, las cuantificaciones se realizaron por duplicado, se analizaron experimentos independientes y con duplicados de las muestras en cada caso.

TABLA VIII. OLIGONUCLEÓTIDOS UTILIZADOS PARA ANÁLISIS DEL RNA MEDIANTE RT-PCR Y Q-RT-PCR

Oligonucleótido ^(a)	Secuencia 5' →3'	Posición ^(b)
Q-LBS7-VS ^(c)	TACCAACACACCATGAAAAAATGC	-
Q-7-RS	TCATTGACAGTTTTAAGATTGAAAGAGTG	28161-28189
Q-2060-VS	TAAAGTTCCATCTTATGACATCGTTTATG	2060-2088
Q-2167-RS	TCTTGTGTGCCAGTTGGTATTTG	12195-12217
SP-VS	GTGAGTGTAGCGTGGCTATATCTCTTC	15-41
7(213)RS	TCTGTAGCAGCAAAATCC	28269-28286
Q1-N-VS	ACTTATGTCCGAGAGACTTTGTACCC	27062-27087
Q1-N-RS	CCAATAACCAATCTGTTGATCCCT	27103-27126
LDVS	CGTGGCTATATCTCTTTTACTTTAACTAG	25-56
7RS	AAAAGTGAATAAATACAGCATGGAGGAA	28080-28108

(a) VS (*virus sense*), sentido viral o polaridad positiva, RS (*reverse sense*), sentido antiviral o polaridad negativa

(b) Posición en el genoma del TGEV. En los oligonucleótidos de sentido inverso se indica la posición de la secuencia complementaria

(c) Oligonucleótido que comprende regiones discontinuas con respecto al genoma del TGEV (Fig. 16)

7. TRANSFECCIÓN DE DNA

7.1. Métodos utilizados para optimizar la transfección de BACs en células de mamífero

Para optimizar la transfección de DNAs de gran tamaño en células de mamífero, se generó la construcción pBAC-NR-TGEV-EGFP de 38 kb, descrita anteriormente, que permite determinar mediante citometría de flujo el porcentaje de células transfectadas, a través de la fluorescencia de la proteína EGFP. Antes de su transfección, el DNA fue precipitado con AcONa/EtOH y resuspendido en agua estéril a una concentración aproximada de 0.5 µg/µl. El proceso de optimización de la transfección se realizó en dos pasos, selección del reactivo de transfección y optimización de las condiciones de transfección con el reactivo seleccionado. Los agentes de transfección que se compararon fueron *Lipofectin* (Invitrogen), *Lipofectamine PlusTM* (Invitrogen), *LipofectamineTM 2000* (Invitrogen), *Effectene* (QIAGEN), *SuperFect* (QIAGEN), *JetPEITM* (Polytransfection) y *Rawa* (Medplant Genetics). En todos los casos se utilizó el protocolo estándar recomendado por el fabricante. El reactivo seleccionado fue

*Lipofectamine*TM 2000 (Invitrogen). En las condiciones optimizadas de transfección las células se sembraron 24 horas antes en medio DMEM con 10% SFT sin antibiótico y se utilizaron al 90-95% de confluencia. Para una placa de 35 mm de diámetro, se utilizó una relación de 5 µg de DNA y 12 µl de *Lipofectamine*TM 2000 y los complejos de transfección se prepararon en un volumen de 500 µl de medio Optimem (GibcoBRL). Previamente a la transfección, las células se lavaron una vez con medio DMEM con 10% SFT sin antibiótico y se repusieron 3/4 partes del volumen de medio de cultivo (1.5 ml). A continuación se añadieron los complejos de transfección y los cultivos se incubaron durante 6 h a 37°C. Transcurrido este tiempo, el medio de transfección se retiró, las células se lavaron, se añadieron 2 ml de medio DMEM con 10% SFT y se incubaron a 37°C hasta su análisis por citometría de flujo.

7.2. Evaluación de la eficiencia de transfección mediante citometría de flujo

Para determinar la eficiencia de transfección de DNA (o de RNA cuando se indique), las células se levantaron con una solución de 0.25% (p/vol) tripsina, 0.02% (p/vol) EDTA y se lavaron con tampón fosfato salino (PBS) utilizando diez volúmenes del medio de cultivo inicial. El medio de lavado se retiró mediante centrifugación a 250g durante 5 min a 4°C y las células se resuspendieron en PBS a una densidad aproximada de 2×10^6 células/ml. El porcentaje de células fluorescentes se determinó utilizando un citómetro *Cell Coulter Epics XL-MCL*.

8. MICROSCOPIA DE FLUORESCENCIA

La infectividad del virus TGEV en células ST y BHK-pAPN se determinó mediante inmunofluorescencia indirecta utilizando el anticuerpo monoclonal de ratón 3BB3 que reconoce la proteína M del TGEV. Las proteínas virales E y N expresadas desde el replicón de Sindbis se analizaron utilizando el anticuerpo monoclonal 3BD10 que reconoce la proteína N o un antisuero policlonal que reconoce la proteína E (S2). La expresión de la proteína EGFP desde el replicón REP 2 del TGEV se analizó mediante microscopía de fluorescencia directa. En todos los casos las células se sembraron sobre cubreobjetos estériles en placas de cultivo de 24 pocillos y se infectaron o transfectaron durante el tiempo especificado. Las células se lavaron tres veces con PBS y se fijaron mediante incubación con paraformaldehído al 4% durante 30 min a temperatura ambiente. Para su análisis mediante inmunofluorescencia indirecta, las células se

lavarón tres veces con PBS y se incubaron durante 1 h 30 min con una solución de bloqueo que contenía saponina con el fin de permeabilizar las células (10% STF, 0.1% saponina, en PBS). La incubación con el anticuerpo primario se realizó en la solución de bloqueo durante 1 h y después de 3 lavados con PBS se realizó la incubación con el anticuerpo secundario conjugado a rodamina durante 30 min. A continuación se realizaron seis lavados de 10-15 min con PBS. Para el montaje de las muestras, los cristales se impregnaron con una solución de glicerol conteniendo el agente de tinción TOPRO 3 (Molecular Probes) para visualizar los núcleos y se observaron utilizando un microscopio *Zeiss Axiovert 200*.

9. METODOLOGÍA PARA LOS ESTUDIOS DE AMPLIFICACIÓN DE MINIGENOMAS POR EL VIRUS COMPLEMENTADOR

9.1 Transcripción *in vitro* de minigenomas derivados del TGEV

La transcripción *in vitro* de los minigenomas se realizó a partir de 1 µg del DNA molde correspondiente linearizado con la enzima *XhoI* en el extremo 3' inmediatamente posterior al terminador de T7, utilizando el reactivo *mMessage-mMachine* (Ambion) que incorpora una estructura cap en el extremo 5' de los transcritos. La reacción se realizó en un volumen final de 20 µl en presencia de H₂O libre de nucleasas, 10 µl de una mezcla de ribonucleótidos trifosfato, conteniendo 15 mM de ATP, CTP y UTP, 3 mM de GTP, y 12 mM de análogo de cap, tampón de transcripción, DNA molde, y 2 µl de una mezcla de enzimas que contiene la RNA polimerasa del fago T7 y el inhibidor de RNAsas *SUPERase-In* (Ambion). En el caso de los minigenomas derivados del RNA DI-C (de 9.7 kb) la concentración de GTP se aumentó añadiendo 3 µl de una solución 20 mM GTP, debido a que este ribonucleótido se encuentra en defecto con respecto a los demás para favorecer la incorporación del cap, lo que reduce la eficiencia de síntesis de transcritos largos. Los componentes de la reacción se mezclaron a temperatura ambiente en el orden indicado y se incubaron a 37°C durante 2 h. Posteriormente, las reacciones se incubaron con 4 U de DNasa I durante 30 min a 37°C para eliminar el DNA molde. El RNA obtenido se purificó utilizando el reactivo *Rneasy kit* (QIAGEN), se cuantificó espectrofotométricamente, se analizó su tamaño e integridad en geles de agarosa desnaturalizantes (1% agarosa-2.2 M formaldehído) y se almacenó en alícuotas a -70°C.

9.2. Transfección de células infectadas con minigenomas

Las células ST o BHK-APN se sembraron en medio DMEM al 10% STF en ausencia de antibiótico en placas de 12 pocillos y se utilizaron al 90-95% de confluencia. Previamente a la transfección de RNA, las células se lavaron una vez y se infectaron con el virus TGEV-PUR-MAD a una mdi de 10, utilizando durante la adsorción medio DMEM 2% STF sin antibiótico. Después de 1 hora y media de adsorción las células se lavaron con medio y se transfectaron con 3 µg ó 1 µg de RNA por cada pocillo, utilizando 6 µl ó 3 µl del reactivo *Lipofectamine 2000* (Invitrogen), cantidades optimizadas para células BHK-APN o ST, respectivamente. En el caso de las células ST, el medio de transfección se retiró después de 4 h, las células se lavaron 3 veces con medio, se añadió medio fresco DMEM al 2% STF y se incubaron a 37°C.

9.3. Análisis del rescate de minigenomas a lo largo de los pases

El RNA de los minigenomas se amplificó mediante pases sucesivos en cultivo. Para ello, la mitad del sobrenadante de los cultivos de células ST o BHK-pAPN infectadas y transfectadas (pase P0) se añadió sobre monocapas confluentes de células ST a las 22-24 hdi y el proceso se repitió durante 4-5 pases sucesivos. En cada pase se recogió el sobrenadante, para determinar el título de virus, y el RNA total, para cuantificar mediante Q-RT-PCR la acumulación de los minigenomas a lo largo de los pases. La cuantificación de los minigenomas se realizó según lo detallado en el apartado 6.3 de esta sección, utilizando oligonucleótidos específicos que flanquearan regiones de secuencia discontinuas respecto al genoma viral. La especificidad de los oligonucleótidos para la detección de los minigenomas se comprobó utilizando RNA procedente de células infectadas y no transfectadas como control negativo.

10. TRANSFECCIÓN Y RESCATE DE VIRUS A PARTIR DE CLONES INFECTIVOS

Para el rescate de virus recombinantes derivados del TGEV que contenían mutaciones puntuales en la posición 637 del genoma, las células BHK-pAPN se crecieron en placas de 35 mm de diámetro y se transfectaron con 5 µg del clon infectivo con la mutación correspondiente en cada caso y 12 µl del reactivo *Lipofectamine 2000* (Invitrogen) según las condiciones optimizadas. A las 6 hdt el medio de transfección se retiró, las células se levantaron con tripsina y se sembraron sobre una monocapa

confluente de células ST. Los sobrenadantes de este cultivo se recogieron a las 24, 36 y 48 horas y se titularon en células ST.

11. ENSAYO DEL PROCESAMIENTO EN TRANS DE LA REPLICASA POR LA PROTEASA VIRAL PLP1

El fragmento N-terminal de la replicasa del TGEV y la proteasa viral PLP1 se expresaron *in vitro* utilizando el reactivo *TNT T7 Quick Coupled Transcription/Translation System* (Promega), que es un sistema de transcripción acoplada a la traducción en lisado de reticulocitos de conejo que contiene la RNA polimerasa del fago T7, siguiendo las instrucciones del fabricante. La síntesis del fragmento N-terminal de la replicasa del TGEV se llevó a cabo utilizando como molde 1 µg de los plásmidos pTGEV-ORF1a(637G) o pTGEV-ORF1a(637A) en presencia de 20 µCi de L-[³⁵S]-metionina (1000 Ci/mmol a 10 mCi/ml; Amersham). La reacción se incubó a 30°C durante 90 min utilizando un volumen de 40 µl de lisado de reticulocitos en un volumen de reacción final de 50 µl. La síntesis del fragmento de la replicasa conteniendo el dominio de la proteasa PLP1 se realizó en ausencia de marca a partir de 1 µg del plásmido pTGEV-PLP1, en presencia de 1 mM metionina y 100 µM ZnCl₂. Una vez transcurridas las reacciones de síntesis del sustrato y la enzima, se añadió cicloheximida a una concentración final de 0.6 µg/µl para bloquear la traducción.

Para analizar la actividad de procesamiento *in vitro* de la proteasa PLP1 sobre los sustratos del fragmento N-terminal de la replicasa con las distintas mutaciones, las mezclas de las reacciones de transcripción/traducción se combinaron utilizando relaciones de enzima-sustrato de 1:2.8, 1:5 y 1:10 (vol/vol) y se incubaron a 30°C durante 16 h. La reacción de procesamiento se paró mediante la adición de tampón de muestra de proteínas 2X [100 mM Tris-HCl pH 6.8, 4% (p/vol) SDS, 0,02% (p/vol) azul de bromofenol, 2% (p/vol) 2-mercaptoetanol, 20% (vol/vol) glicerol]. Dos tercios del total de las reacciones se resolvieron mediante electroforesis en geles de poliacrilamida al 10% en presencia de SDS (SDS-PAGE) según el procedimiento descrito por Laemmli (Laemmli, 1970). Los geles de proteínas se fijaron mediante incubación con una solución de metanol-ácido acético-agua (5:1:4 [vol/vol/vol]) durante 30 min. Posteriormente, se incubaron en presencia del reactivo *Amplify NAMP199* (Amersham) durante 30 min, se secaron y se revelaron mediante autorradiografía

utilizando películas Kodak X-OMAT AR film a -70°C durante tiempos comprendidos entre 1 y 36 h.

12. CARACTERIZACIÓN DE REPLICONES DERIVADOS DEL GENOMA DEL TGEV

12.1. Análisis de la actividad de los replicones mediante RT-PCR y Q-RT-PCR

Para analizar la actividad de los replicones del TGEV, el RNA total intracelular se aisló a los 1.5, 5 o 10 días después de la transfección utilizando el reactivo *RNeasy Mini Kit* (QIAGEN). La actividad de los replicones se determinó mediante el análisis del RNA mensajero del gen 7 mediante RT-PCR convencional utilizando los oligonucleótidos SP-VS y 7(213)RS (Tabla VIII), y mediante Q-RT-PCR utilizando los oligonucleótidos LDVS y 7RS (Tabla VIII), siguiendo el protocolo descrito para la cuantificación de los minigenomas en el apartado 6.3 de esta sección.

12.2. Determinación de la actividad β -glucuronidasa

El análisis de la actividad en *trans* del replicón mínimo teórico TGEV-REP se realizó cuantificando la expresión del gen de la β -glucuronidasa a partir del minigenoma M39-GUS. Para ello la actividad enzimática se determinó por quimioluminiscencia utilizando el sistema *GUS-Light kit* (Tropix), siguiendo las instrucciones del fabricante. Las células se sembraron en placas de 35 mm de diámetro y se cotransfectaron con el replicón TGEV-REP (5 μ g) y el minigenoma M39-GUS (1 μ g), utilizando 15 μ l de *Lipofectin* (Invitrogen). Las células se recogieron a los tiempos indicados, se lavaron con PBS y se resuspendieron en 200 μ l de tampón de lisis (100 mM tampón fosfato sódico pH 7.8, 0.2% Tritón X-100, 1 mM DTT). Los extractos celulares o diluciones seriadas de los mismos (6.7 μ l), se incubaron en tubos de luminómetro a temperatura ambiente durante 1 h, en presencia de 60 μ l del sustrato quimioluminiscente *Glucuron* (Bronstein y col., 1994) diluido 1:100 en tampón de reacción (100 mM tampón fosfato sódico pH 7.0, 10 mM EDTA). Los tubos con la mezcla de reacción se colocaron en el luminómetro y la reacción se potenció mediante la inyección de 100 μ l de un acelerador de la emisión de luz *Light Emisión Accelerator* (Bronstein y col., 1996). La emisión de luz durante los 10 primeros segundos de potenciación de la reacción se midió en un luminómetro *Lumat LB 9501* (Berthold Systems) y se expresó como unidades luminométricas relativas (RLUs).

12.3. Análisis del efecto citopático inducido por los replicones en diferentes líneas celulares

La actividad de los replicones derivados del TGEV se analizó en células humanas 293T y HeLa, en células de hámster BHK-pAPN y en células porcinas ST. La construcción pBAC-TGEV-NR-EGFP se transfectó en paralelo para determinar el porcentaje de células transfectadas mediante citometría de flujo. Los replicones fueron transfectados en los distintos tipos celulares utilizando el reactivo *Lipofectamine 2000*, las células se incubaron a 37°C y posteriormente el RNA total intracelular se recogió entre las 12-40 hdt para analizar la actividad de los replicones mediante RT-PCR. En los casos en que la viabilidad del cultivo lo permitiera, las células se pasaron realizando una dilución de una sexta parte del cultivo. En cada pase se analizó el porcentaje de células transfectadas con la construcción control pBAC-TGEV-NR-EGFP mediante citometría de flujo, y se extrajo el RNA de las células transfectadas con los replicones, para analizar su actividad mediante RT-PCR.

12.4. Generación de células BHK-pAPN transfectadas con el replicón del virus Sindbis expresando diferentes mutantes del gen N

La generación de células BHK-pAPN expresando la proteína N viral se realizó de la forma descrita anteriormente para la proteína E (Ortego y col., 2002) con ligeras modificaciones. Las células BHK-pAPN se sembraron en medio DMEM con 10% SFT sin antibiótico en placas de 10 cm de diámetro y se transfectaron semiconfluentes con la construcción pSINrep21-N utilizando 12 µg de plásmido y 60 µl del reactivo *Lipofectamine 2000*. Las células se incubaron a 37°C durante 6 h en presencia de los complejos de transfección, se lavaron y se añadió medio DMEM 10% SFT sin antibiótico. A las 24 hdt el medio de cultivo se cambió por el de selección, conteniendo 5 µg/ml de puromicina, y se incubaron durante tres días hasta alcanzar la confluencia, realizando un cambio de medio diario. Una vez seleccionadas las células que expresaban constitutivamente el replicón pSINrep21-N, estas se transfectaron con los diferentes replicones del TGEV. Para ello las células se sembraron en placas de 35 mm de diámetro con medio DMEM al 10% SFT sin antibióticos y la transfección se realizó utilizando 5 µg de DNA y 12 µl de *Lipofectamine 2000*. Las células se incubaron a 37°C durante 4 h en presencia de los complejos de transfección, se lavaron y se añadió medio fresco DMEM 10% SFT sin antibiótico. A las 24 hdt se determinó el porcentaje de

células transfectadas y se recogió el RNA para analizar la actividad de los replicones mediante RT-PCR.

Para analizar la influencia de la proteína N o de su RNA no traducible sobre la actividad de los replicones, se generaron células BHK-pAPN que expresaban desde el replicón del virus Sindbis el RNA de la proteína N, con mutaciones en los codones de metionina de este gen para impedir su traducción. Para ello las células BHK-pAPN se transfectaron con las construcciones pSINrepN1, pSINrepN2 o pSINrepN3 descritas en el apartado 5.6 de esta sección, utilizando *Lipofectamine 2000*. Estas células, junto con las que expresaban el mensajero sin mutar de la proteína N, se transfectaron con la construcción pBAC-TGEV-REP3 y 24 h después se determinó la actividad del replicón mediante Q-RT-PCR. En paralelo se cuantificó la cantidad de RNA mensajero transcrito desde el replicón de Sindbis en las diferentes células mediante Q-RT-PCR utilizando los oligonucleótidos Q1-N-VS y Q1-N-RS (Tabla VIII).

13. MÉTODOS PARA EL AISLAMIENTO Y CARACTERIZACIÓN DE PROTEÍNAS QUE INTERACCIONAN CON LOS EXTREMOS DEL VIRUS

13.1. Fraccionamiento subcelular y preparación de extractos de proteínas

Los extractos de proteínas citoplasmáticas se prepararon a partir de monocapas confluentes de células ST o BHK-pAPN que se infectaron en los casos indicados con el virus TGEV a una mdi de 10 y se recogieron a las 7 hdi. Las células se lavaron dos veces con 25 ml de PBS previamente enfriado a 4°C, se levantaron con una espátula, se centrifugaron a 250g durante 10 min a 4°C, se resuspendieron en PBS a una concentración estimada de 4×10^7 células/ml y se congelaron a -70°C hasta su procesamiento. Los extractos de proteínas se prepararon frescos para cada experimento partiendo de un precipitado de 4×10^7 células congeladas resuspendidas en PBS. Para ello las células se centrifugaron a 250g durante 10 min a 4°C, se resuspendieron en 400 µl de tampón de extracción (10 mM HEPES pH 7.9, 10 mM KCl, 0.1 mM EDTA, 1 mM DTT y cóctel de inhibidores de proteasas *Complete* de Roche) y se lisaron a 4°C durante 15 min. A continuación se añadieron 25 µl de NP-40 al 10% y después de agitar al vortex 10 seg, los núcleos se eliminaron por centrifugación a 13000g durante 30 seg. El sobrenadante se congeló a -70°C en presencia de glicerol al 10% hasta su uso. Todas las soluciones no comerciales para trabajar con proteínas o RNA descritas en esta sección y sucesivas se filtraron utilizando filtros *Millex* (Millipore) de 0.22 µm.

13.2. Síntesis y purificación de RNA biotinado

El RNA utilizado para los ensayos de interacción con proteínas fue sintetizado mediante transcripción *in vitro* a partir de los moldes de transcripción pT7-3'TGEV-*Hind*III-MBN, pM33L-SP6-5'TGEV-*Eco*RV, pM33L-SP6-5'TGEV-*Bgl*II y PCR-SP6-5'TGEV-0.5kb cuya construcción se ha descrito en el apartado 5.7 de esta sección. La síntesis se realizó utilizando el reactivo *SP6/T7 Transcription kit* (Roche) según las instrucciones del fabricante, con la polimerasa adecuada según el promotor de transcripción en cada caso. Para el marcaje del RNA se añadió a la reacción Biotin-14-CTP (GibcoBRL) en una relación 1:5.25 con respecto al CTP no marcado, quedando a una concentración final de 0.16 mM. Las reacciones se incubaron durante 2 h a 37°C y a continuación se trataron con 10 U de DNAsa I durante 30 min para eliminar el DNA molde. Los transcritos se purificaron utilizando el reactivo *RNeasy kit* de QIAGEN, se analizaron en geles desnaturalizantes de 1% agarosa-2.2 M formaldehído, se cuantificaron espectrofotométricamente y se almacenaron en alícuotas a -70°C.

13.3. Cromatografía de afinidad de RNA

La captura de proteínas por afinidad a RNA se realizó mediante cromatografía de afinidad utilizando como matriz *Streptavidin Sepharose High Performance* (Amersham) y una solución de unión y lavado (BW) que contenía 50 mM HEPES pH 7.9, 150 mM KCl, 5% glicerol y 0.01% NP-40. La cromatografía se realizó en tubos de PCR, realizando las incubaciones a 4°C en un agitador orbital, excepto que se indique de otro modo, y sedimentando la resina por centrifugación a 440g durante 1 min utilizando un rotor basculante. Los pasos de lavado se realizaron con 200 µl de solución BW durante 15 min. Antes de la cromatografía, el extracto de proteínas (1-5 µg/µl) se preaclará tres veces utilizando un volumen de 20 µl de resina cada vez (volúmenes referidos a la resina sedimentada) por cada 100 µl de extracto. Los pasos de preaclaredo se realizaron durante cuatro horas, después de las cuales se eliminó la resina y el extracto de proteínas se transfirió a un tubo nuevo con resina, repitiendo el proceso tres veces. En paralelo al último paso de preaclaredo, el RNA biotinado (1-10 µg de RNA según se indique) se inmovilizó mediante su incubación con 20 µl de resina en presencia de un tampón de alta concentración de sal conteniendo 5 mM HEPES pH 7.9, 0.5 mM EDTA, 1 M KCl. Posteriormente, la resina con el RNA inmovilizado se lavó tres veces con la solución BW para eliminar el exceso de sal y se incubó con los extractos de proteínas

preaclarados. En cada tubo de PCR conteniendo 20 µl de resina con el RNA se añadieron 100 µl de solución BW y 100 µl de extracto de proteínas y se incubaron durante 14-16 h. Posteriormente, los complejos se aislaron por centrifugación, se lavaron tres veces y las proteínas unidas al RNA se disociaron mediante resuspensión de la resina con 20 µl de tampón de muestra 1X *NuPage* (Invitrogen) conteniendo 50 mM DTT e incubación a 65°C durante 10 min en un bloque termociclador. La resina se centrifugó, se recuperó el sobrenadante y se realizó una segunda elución con 12 µl de tampón de muestra, que se mezcló con el primer sobrenadante y se almacenó a -70 °C hasta su análisis.

13.4. Preparación de muestras para su análisis mediante espectrometría de masas

Las proteínas eluidas de la cromatografía de afinidad de RNA se separaron mediante electroforesis desnaturalizante en geles de poliacrilamida comerciales *NuPAGE 4-12% Bis-Tris gel 1.0 mm* (Invitrogen) utilizando el tampón *1X NuPAGE MOPS SDS Running buffer* (Invitrogen) como solución electrolito. En la cubeta superior se añadió 0.5 ml de antioxidante *NuPAGE antioxidant* (Invitrogen) para evitar que las proteínas previamente reducidas por el DTT se oxidaran durante la electroforesis, que se realizó a 100 V durante 3 h aproximadamente para una buena definición de las bandas. Posteriormente, el gel se lavó tres veces durante 5 min con 100 ml de agua desionizada y se tiñó con azul de coomasie *Simply Blue Safe Stain* (Invitrogen) según las indicaciones del proveedor. Después de la tinción se tomaron imágenes de los geles utilizando un scanner *UMAX Astra 4000U*.

Las bandas de proteína se recortaron del gel con un bisturí tomando el mínimo volumen de gel, se conservaron a 4 °C en agua desionizada y se digirieron con tripsina, utilizando un digestor de proteínas automático *Proteineer DP* (Bruker Daltonics, Bremen, Alemania). La elución de las muestras para su análisis mediante espectrometría de masas MALDI-TOF se realizó con ácido trifluoroacético al 0.1 %. Alternativamente, la elución de las muestras para su análisis mediante espectrometría de masas ESI-QUAD-TOF se realizó en dos pasos con acetonitrilo al 50% diluido en agua. En este último caso, los péptidos extraídos se concentraron hasta un volumen de 2-3 µl en un concentrador de vacío *SpeedVac* y se almacenaron a -20°C hasta su análisis.

13.5. Espectrometría de masas de desorción/ionización por laser asistida por matriz y acoplada a un detector de tiempo de vuelo (MALDI-TOF)

Para la obtención de las masas de los péptidos trípticos, 0.5 µl de la solución matriz (1 mg/ml ácido 4-hidroxi-cinámico) se depositó manualmente sobre un portamuestras *AnchorChipTM MALDI* de 600 µm y se dejó secar a temperatura ambiente. A continuación se depositó el mismo volumen de la mezcla de péptidos trípticos y se dejó secar de nuevo a temperatura ambiente. Una vez cocrystalizado el analito con la matriz, la muestra se ionizó mediante su irradiación con un láser utilizando un espectrofotómetro de masas *MALDI-TOF ReflexTM IV Bruker* (Bruker-Franzen Analytic GMBH, Bremen, Alemania) equipado con una fuente *SCOUTTM* en modo reflector de iones positivos utilizando extracción retardada y el programa *FlexControl v.2.1* en modo automático de adquisición. El voltaje de aceleración de iones fue de 25 kV.

13.6. Espectrometría de masas mediante ionización por electrospray acoplada a un analizador de masas de tipo cuádrupolo y un detector de tiempo de vuelo (ESI-QUAD-TOF)

Las muestras de proteína digeridas y almacenadas a -20°C fueron resuspendidas en 6 µl de ácido fórmico al 0.1% e inyectadas para su análisis en un espectrómetro de masas *Q-TOF Micromass* (Manchester, Reino Unido) acoplado a un sistema de cromatografía líquida capilar de alta resolución *CapLC HPLC*. La muestra fue retenida y desalada en una columna *C18 Pepmap trapping column* (Dionex Ltd, Camberley, Surrey, Reino Unido) de 300 micras de diámetro interno y 5 mm de longitud. Los péptidos fueron eluidos y resueltos en una columna analítica *Atlantis C18* (Waters, Milford, Massachussets, EEUU) de 75 micras de diámetro interno y 5 cm de longitud aplicando un gradiente de acetonitrilo del 5-90% diluido en agua conteniendo 0,1% de ácido fórmico. La elución con el gradiente se realizó durante 30 min con un flujo de 250 nl por minuto. El análisis con el espectrómetro de masas *Q-TOF Micromass* fue realizado en modo de adquisición dependiente de datos en el que se seleccionaron iones con carga múltiple (+2 y +3) para su fragmentación inducida por colisión en presencia de gas argón. Los datos fueron analizados utilizando el programa *MassLynx v.4.0* (Micromass, Manchester, Reino Unido).

13.7. Herramientas informáticas para la identificación de proteínas

Los datos adquiridos mediante espectrometría de masas MALDI-TOF fueron procesados con el programa *FlexAnalysis v.2.4* para el calibrado, suavizado y etiquetado de péptidos. A continuación se transfirieron a través del programa *MS BioToolsTM* para realizar búsquedas automatizadas frente a la base de datos del NCBIInr utilizando la aplicación *Mascot* (Matrix Science, Londres, Reino Unido) desde el sitio web <http://www.matrixscience.com>. En los casos pertinentes se introdujeron restricciones de taxonomía para la búsqueda de candidatos, seleccionando la base de datos de mamíferos, roedores o virus. Se tuvo en cuenta como modificación fija la carbamidometilación de las cisteínas y como modificación variable la oxidación de las metioninas. La masa molecular para la búsqueda de proteínas candidato no se restringió a ningún valor y los valores de las masas de los péptidos trípticos seleccionados fueron las correspondientes a especies monoisotópicas y monoprotonadas. Se permitió la ausencia de un corte por la tripsina y se utilizó una tolerancia de 0.1 Da en la búsqueda de los péptidos.

A partir de los datos obtenidos mediante espectrometría de masas ESI-QUAD-TOF, la búsqueda de proteínas candidato se realizó de forma similar utilizando el modo *Mascot MS/MS ion search* desde el sitio web <http://www.matrixscience.com>, con la modificación de algunos parámetros. Se utilizó la base de datos del NCBIInr sin restricciones taxonómicas. Las masas de los péptidos trípticos seleccionadas para la búsqueda fueron las correspondientes a especies monoisotópicas y se definió una tolerancia de 0.2 Da en la búsqueda de iones precursores y sus productos de fragmentación.

Para cada identificación, se comprobó que la masa molecular de la proteína asignada correspondiera a la estimada por movilidad de la proteína recortada del gel.

14. MÉTODOS PARA EL ANÁLISIS DE LA INTERACCIÓN DE LOS EXTREMOS DEL GENOMA VIRAL

14.1 Síntesis y purificación de RNA marcado con digoxigenina o biotina

El RNA utilizado para los ensayos de interacción entre los extremos del genoma del TGEV fue sintetizado mediante transcripción *in vitro* a partir de los moldes de transcripción pT7-3'TGEV-*Hind*III-MBN y pM33L-SP6-5'TGEV-*Eco*RV cuya construcción se ha descrito en el apartado 5.7 de esta sección. La síntesis y purificación

del RNA se realizó como se ha descrito anteriormente para la síntesis de RNA biotinado en el apartado 13.2 de esta sección, utilizando para el marcaje alternativamente Biotin-14-CTP (GibcoBRL) o DIG-11-UTP (Roche) que se añadieron a la reacción de transcripción en una relación 1:5.25 con respecto al ribonucleótido correspondiente no marcado, quedando a una concentración final de 0.16 mM.

El marcaje del RNA se comprobó mediante ensayos de *dot-blot*. Para ello los RNAs se cuantificaron espectrofotométricamente, se prepararon diluciones seriadas de orden diez de las que se tomó un volumen de 2 µl que se adsorbió directamente sobre una membrana de nailon cargada positivamente *BRIGHTSTAR-PLUS Positively Charged Nylon Membrane* (Ambion). A continuación los RNAs se inmovilizaron en la membrana mediante tratamiento con luz UV aplicando dos pulsos de energía de 700 x 100 µJ/cm² utilizando un equipo *CL-1000 Ultraviolet crosslinker* (UVP, San Gabriel, CA). Para la detección del RNA marcado con digoxigenina se utilizó un anticuerpo anti-digoxigenina acoplado a fosfatasa alcalina (Roche) a una dilución 1/30000 y para la detección del RNA marcado con biotina se utilizó el conjugado de estreptavidina con fosfatasa alcalina incluido en el sistema de revelado. Para el revelado de las membranas se utilizó el sistema *BrightStar BioDetect* (Ambion) según las instrucciones del fabricante.

14.2. Ensayo de coprecipitación de RNA mediada por proteínas

Para el ensayo de coprecipitación de RNA se utilizaron fragmentos de RNA de los extremos del virus sintetizados *in vitro* y marcados de forma diferencial con digoxigenina o biotina. En primer lugar se realizó una cromatografía de afinidad de RNA utilizando 1 µg del RNA biotinado y extracto citoplasmático de proteínas según el protocolo detallado en el apartado 13.3 de esta sección, con las modificaciones descritas a continuación. Una vez capturadas las proteínas de unión al RNA, y después de un primer lavado durante 15 min con solución BW, se añadieron 250 ng del otro fragmento de RNA marcado con digoxigenina diluido en 200 µl de la solución BW en presencia de 40 U de inhibidor de RNasas *Superase-In* (Ambion). La incubación en presencia del segundo RNA se realizó durante 30 min a 4 °C en un agitador orbital. En los casos que se indique se añadió un exceso de RNA ribosomal 16S-23S de *E.coli* (Roche) o RNA no marcado como competidor durante esta incubación. A continuación se realizaron tres lavados con 200 µl de tampón BW, incubando el primero durante 15 min. Se añadió 1

µg de tRNA de levadura (Sigma) y 40 U de inhibidor de RNAsas a los complejos recuperados por centrifugación, en un volumen de 10 µl de la solución BW conteniendo 1 mM de EDTA y se agitó manualmente para resuspender la resina de estreptavidina-sefarosa sedimentada. Seguidamente se añadieron 1.5 µg de proteasa K (Ambion), diluida en 10 µl de la misma solución, y se incubaron a 37°C durante 10 min en un bloque termociclador. Para analizar el RNA coprecipitado, se añadieron 20 µl de tampón de muestra de RNA desnaturalizante *RNA loading buffer II* (Ambion), la resina se resuspendió mediante agitación manual y se incubó a 80°C durante 2 min en un bloque termociclador. Posteriormente, la resina se eliminó por centrifugación y el RNA contenido en el sobrenadante se almacenó a -70°C hasta su análisis.

14.3. Electroforesis, transferencia, fijación y detección del RNA en membranas de nailon

El RNA coprecipitado se separó en geles desnaturalizantes de poliacrilamida al 4% y 8 M urea (Sambrook y Russell, 2001) realizando la electroforesis a 150 V y 4°C durante 1 h. La transferencia de RNAs a membranas de nailon se realizó a 200 mA durante 1 h utilizando las membranas *BRIGHTSTAR-PLUS Positively Charged Nylon Membrane* (Ambion) y 0.5X TBE (90 mM Tris-Borato, 2mM EDTA, pH 8.0) como solución electrolito. A continuación los RNAs se fijaron a la membrana mediante tratamiento con luz UV aplicando dos pulsos de energía de 700 x 100 µJ/cm² utilizando un equipo *CL-1000 Ultraviolet crosslinker* (UVP, San Gabriel, CA).

Para visualizar el RNA coprecipitado, las membranas se revelaron para la detección específica del RNA marcado con digoxigenina, utilizando las soluciones de bloqueo, lavado y revelado del sistema *BrightStar BioDetect* (Ambion) según el protocolo del fabricante y el anticuerpo anti-digoxigenina acoplado a fosfatasa alcalina (Roche) a una dilución 1/30000. Cuando se utilizó este anticuerpo el protocolo de revelado se modificó añadiendo tres lavados adicionales de 15 min. El marcaje de los RNAs con biotina se comprobó en los casos necesarios utilizando los mismos reactivos y el conjugado de estreptavidina con fosfatasa alcalina incluido en el sistema de revelado.

VI. RESULTADOS

1. EFECTO DIFERENCIAL DE UNA MUTACIÓN PUNTUAL EN EL GEN DE LA REPLICASA SOBRE LA REPLICACIÓN DEL VIRUS TGEV Y LA AMPLIFICACIÓN DE MINIGENOMAS

1.1. REVERSIÓN DE UN MARCADOR GENÉTICO EN EL GENOMA DEL VIRUS RECOMBINANTE RESCATADO A PARTIR DEL CLON INFECTIVO DEL TGEV

El clon infectivo del virus TGEV se construyó en un BAC tomando como base el RNA defectivo DI-C derivado del aislado respiratorio PUR46-MAD (Fig. 10) (Almazán y col., 2000). Durante su construcción se introdujeron como marcadores genéticos dos mutaciones silenciosas en las posiciones 6752 y 18997 del genoma y una mutación en la posición 637 que daba lugar a un cambio de aminoácido (Gly por Val), debido a que esta sustitución se encontraba en el RNA defectivo DI-C y sus minigenomas derivados como el M33L, que son replicados eficientemente por el virus complementador (Fig. 10).

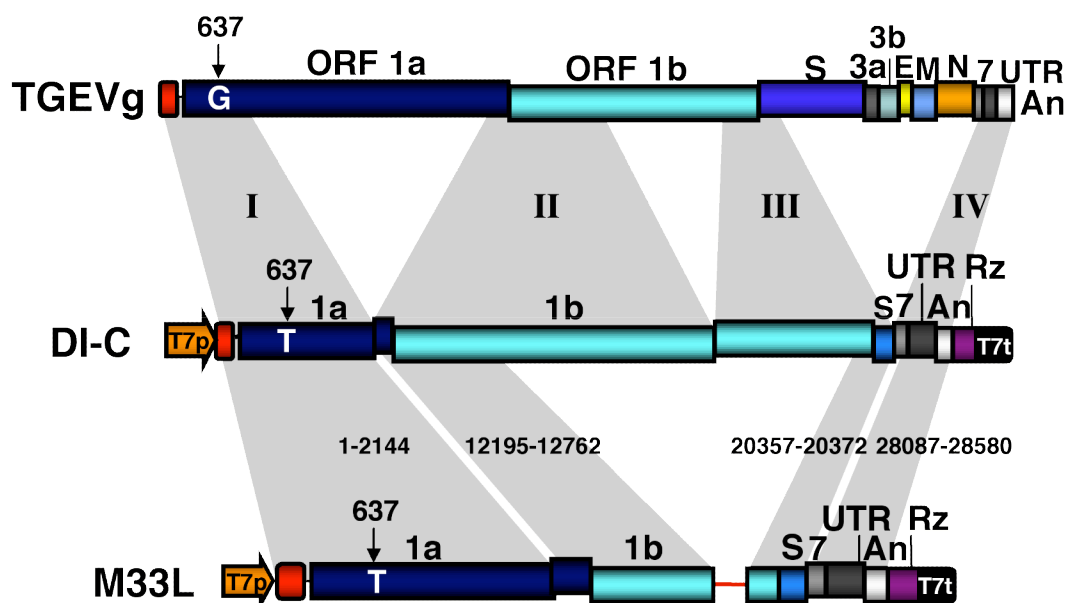


Figura 10. Estructura genética del TGEV y de RNAs defectivos derivados del mismo. Las barras representan el genoma del virus TGEV y los cDNAs del RNA defectivo DI-C y del minigenoma derivado M33L. Ambos RNAs defectivos están formados por cuatro regiones discontinuas del genoma parental, numeradas del I al IV en el RNA DI-C. Las posiciones de los nucleótidos con respecto al genoma se indican en el minigenoma M33L. La correspondencia de secuencia entre las distintas construcciones se indica mediante polígonos sombreados. Los genes virales se indican sobre las barras. An, poliA; T7p, promotor del fago T7; T7t, terminador de la transcripción de la polimerasa de T7; HDV Rz, secuencia de la ribozima del virus de la hepatitis delta.

Cuando se analizó la secuencia de los marcadores genéticos del virus recombinante rescatado a partir de transfecciones independientes del clon infectivo se observó una reversión del marcador genético en la posición 637 de una T a una G, que es la secuencia viral consenso. Sin embargo, los otros marcadores genéticos se mantuvieron estables, indicando una fuerte presión de selección sobre la posición 637 del genoma (Fig. 11). Esta observación sugería que el nucleótido de la posición 637 afectaba a un motivo del RNA o residuo de la replicasa que estaba desempeñando un papel esencial en el ciclo infectivo del virus. Con el fin de estudiar la importancia de la secuencia en la posición 637 para la replicación del virus así como para la amplificación *in trans* de los minigenomas derivados del TGEV, se realizó un análisis mutacional en esta posición utilizando para ello el clon infectivo del TGEV y los cDNAs de los minigenomas DI-C y M33L.

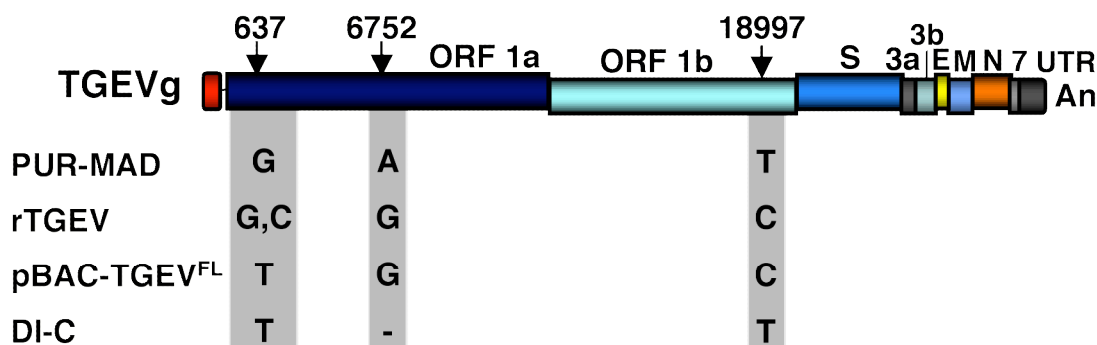


Figura 11. Análisis de los marcadores genéticos en los virus rTGEV rescatados a partir del cDNA. Esquema de la estructura genética del clon infectivo del TGEV, en el que se indica mediante flechas la posición de los marcadores genéticos introducidos. La secuencia de los marcadores se muestra para el virus parental (PUR-MAD), los virus recombinantes rescatados (rTGEV), el clon infectivo (pBAC-TGEV^{FL}) y el RNA defectivo a partir del cual fue reconstruido (DI-C). Las letras y números sobre las barras indican los genes virales. UTR, región 3' no traducida; An, poliA.

1.2. UNA MUTACIÓN PUNTUAL EN EL GEN DE LA REPLICASA AFECTA DRÁSTICAMENTE AL RESCATE DE VIRUS A PARTIR DEL CLON INFECTIVO DEL TGEV

Con el fin de analizar el papel del nucleótido 637 en la replicación del virus, se construyeron clones cDNA del TGEV sustituyendo en cada caso la secuencia viral consenso G por C, T, o A, que dan lugar a un cambio de aminoácido en la posición 108 de las poliproteínas de la replicasa (pp1a y pp1ab) de Gly a Ala, Val o Asp, respectivamente (Fig. 12A). Los mutantes generados se transfectaron en células BHK-

pAPN, y 6 hdt las células se levantaron con tripsina y se sembraron sobre una monocapa de células ST, para amplificar la producción de virus.

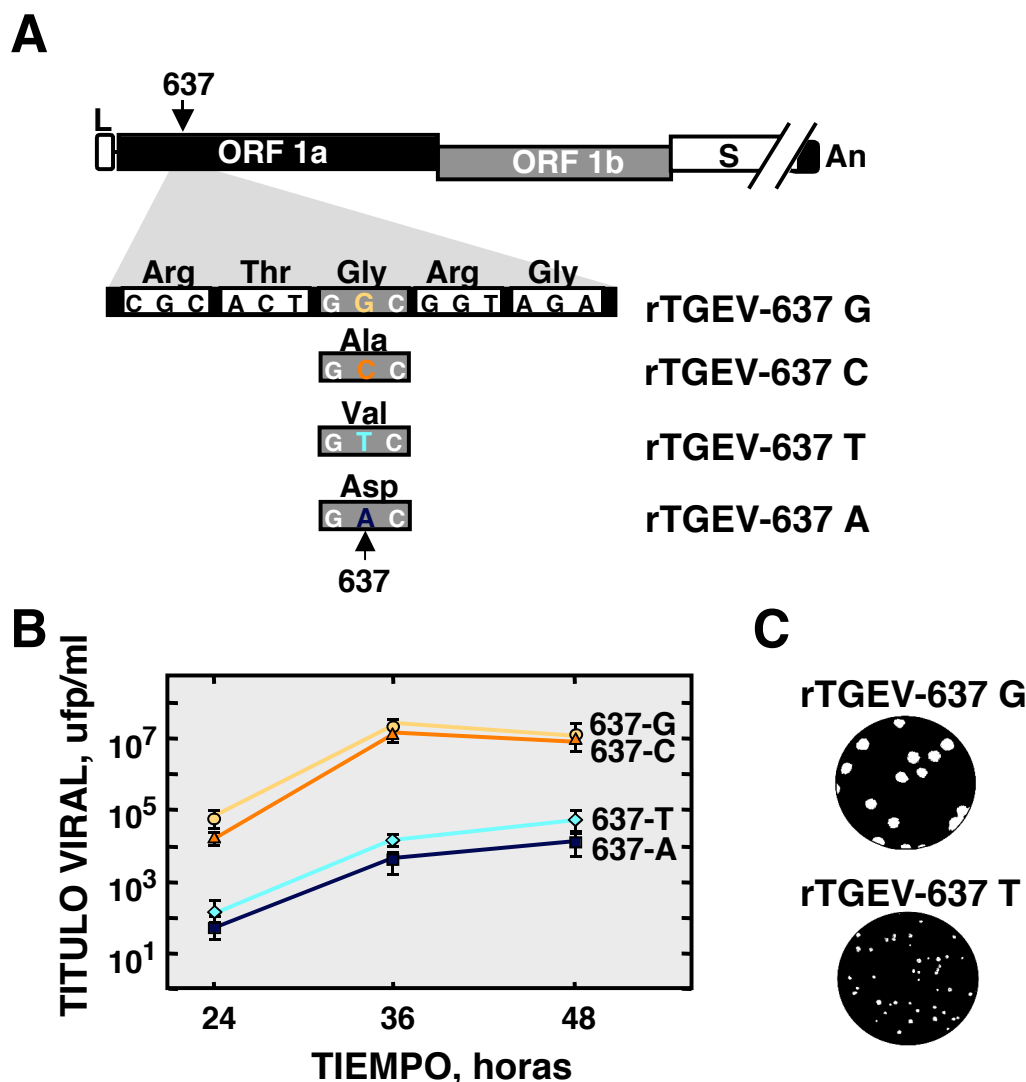


Figura 12. Fenotipo de los virus recombinantes con mutaciones puntuales en la posición 637. (A) Esquema de la organización del cDNA del TGEV con las cuatro posibles mutaciones del nucleótido 637, indicando el cambio de aminoácido al que dan lugar. (B) Rescate de virus. Los cDNAs del TGEV con las mutaciones en la posición 637 se transfectaron sobre células BHK-pAPN. Los sobrenadantes de estos cultivos se recogieron a los tiempos indicados y se titularon en células ST. (C) Morfología de placa de los virus recombinantes. Los fenotipos de los virus rTGEV-637C y rTGEV-637A fueron similares a los mostrados para los virus rTGEV-637G y rTGEV-637T, respectivamente.

Los sobrenadantes de estos cultivos se recogieron a las 24, 36 y 48 hdt y el rescate de virus se analizó mediante titulación de estos sobrenadantes en células ST (Fig. 12B).

Se observó una reducción de tres unidades logarítmicas en el rescate de virus recombinantes a partir de los cDNAs que contenían una T o una A en la posición 637, en correlación con una reducción en el tamaño de placa de lisis con respecto al virus

salvaje (Fig. 12C). Por el contrario, los títulos de virus rescatados a partir del mutante que contenía una C en la posición 637 fueron similares a los obtenidos a partir del cDNA con la secuencia viral consenso (637G) y no se observaron diferencias significativas en el tamaño de placa de lisis. El fenotipo de los virus con las mutaciones 637A y 637T fue inestable, y se observaron revertientes con morfología de placa grande con una G o una C en la posición 637 desde tiempos tempranos después de la transfección. Estos revertientes constituyeron la población viral dominante después del primer pase de virus (resultados no mostrados). A partir de estas observaciones se estableció una correlación entre las mutaciones 637A y 637T, que daban lugar a los cambios más drásticos de aminoácido y aquellos virus recombinantes afectados en replicación, sugiriendo que el efecto de las mutaciones puntuales en la posición 637 del genoma era probablemente a nivel de proteína.

1.3. ESTUDIO DE LA REPLICACIÓN DE MINIGENOMAS DERIVADOS DEL TGEV

Con el objetivo de generar un sistema para evaluar la replicación viral de forma independiente de la encapsidación se utilizó el minigenoma M33L (3.3 kb), el más pequeño que es rescatado eficientemente, para poner a punto un ensayo que permitiera evaluar su replicación en el pase P0. Este minigenoma está modificado a partir del M33 mediante la inserción de una secuencia espaciadora o *linker* que no afecta a su rescate eficiente (Izeta, 2000). A diferencia con el genoma viral, los minigenomas sintéticos M33 y M33L contienen una T en la posición 637, que estaba presente en el RNA defectivo natural DI-C del que derivan por delección. Para realizar los estudios de replicación con la secuencia correcta del virus, se corrigió la mutación que presentaba el minigenoma M33L en la posición 637, sustituyendo la T del minigenoma por una G, que es la secuencia viral consenso.

Hasta la fecha, el abordaje de los estudios de replicación en el pase P0 con minigenomas del TGEV se había visto dificultado por la baja eficiencia de transfección de los minigenomas en las células ST y la necesidad de un sistema de amplificación para su detección. Para resolver estos problemas se decidió optimizar la transfección del RNA de los minigenomas y evaluar su amplificación en el pase P0 mediante Q-RT-PCR.

Para mejorar la eficiencia de la transfección de los minigenomas se utilizaron células BHK-pAPN que expresan el receptor para el TGEV y que son transfectadas con mayor

eficiencia que las células porcinas ST. Se analizó la susceptibilidad de las células BHK-pAPN a la infección por el TGEV y la producción de virus en comparación con las células ST. Ambos tipos celulares se infectaron a una mdi de 10 y a las 14 hdi se analizó el porcentaje de células infectadas mediante inmunofluorescencia utilizando el anticuerpo monoclonal 3BB3 que reconoce la proteína M del TGEV. El porcentaje de células infectadas en los dos tipos celulares fue próximo al 100% (Fig. 13A). A continuación se realizó una cinética de producción viral. Para ello, monocapas

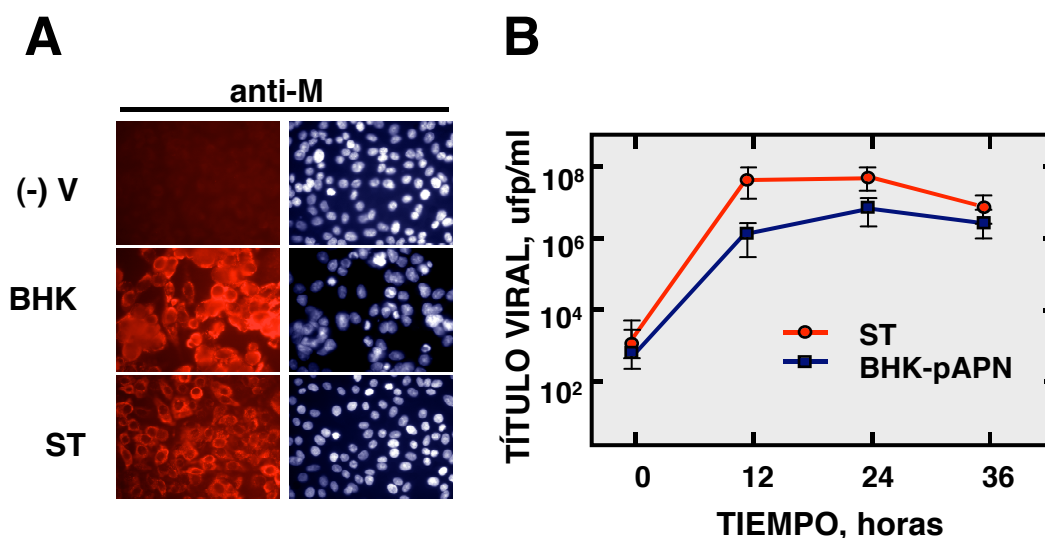


Figura 13. Infectividad y producción viral en células ST y BHK-pAPN. (A) Inmunofluorescencia. La infectividad de las células ST y BHK-pAPN por el TGEV se analizó mediante visualización de la proteína M viral por inmunofluorescencia indirecta utilizando el anticuerpo monoclonal específico 3BB3; (panel izquierdo). Los núcleos de las células se visualizaron mediante tinción con el reactivo TOPRO 3 (panel derecho). BHK, células BHK-pAPN; (-)V, células no infectadas. (B) Cinética de producción de virus. Las células ST y BHK-pAPN se infectaron con el TGEV a una mdi de 10. El sobrenadante de estos cultivos se recogió a los tiempos indicados y se tituló en células ST.

confluentes de células ST y BHK-pAPN se infectaron con el TGEV a una mdi de 10 y el sobrenadante se recogió a los tiempos indicados para su titulación en células ST (Fig. 13B). Se observó una reducción del título viral máximo de una unidad logarítmica y una cinética más retrasada en el caso de las células BHK-pAPN ST en comparación con las ST.

Una vez comprobada la infectividad de las células BHK-pAPN, se optimizó la transfección de RNA para evaluar el rescate de minigenomas en estas células. Para ello, se construyó el minigenoma M33L-EGFP que está clonado bajo el promotor del fago T7 y expresa la proteína EGFP bajo el IRES del virus de la encefalomiocarditis. Este

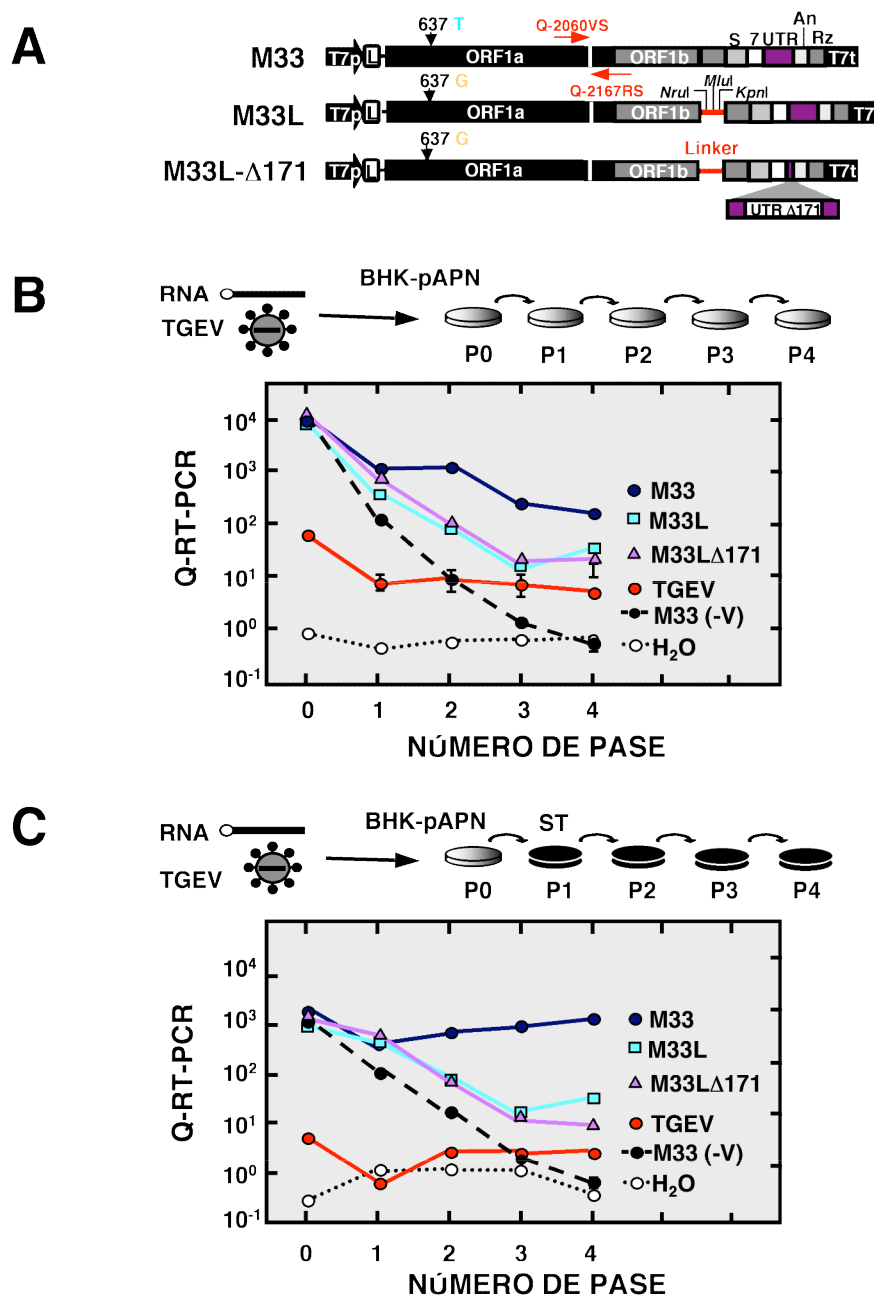


Figura 14. Rescate de minigenomas en células BHK-pAPN y ST. (A) Esquema de los minigenomas M33, M33L y M33L-Δ171, indicando los genes virales y secuencias flanqueantes. T7p y T7t, promotor y terminador de T7, respectivamente; UTR, región 3' no traducida; An, poliA; Rz, secuencia de la ribozima del HDV. Los oligonucleótidos utilizados para el análisis por Q-RT-PCR de los minigenomas se indican con flechas rojas sobre el M33. La secuencia *Linker* se representa por una línea roja. (B) Rescate de los minigenomas en células BHK-pAPN. Los minigenomas se transfectaron en células BHK-pAPN infectadas. A intervalos de 22-24 h los sobrenadantes se pasaron sobre células BHK-pAPN y en cada pase se recogió el RNA (esquema superior). La acumulación de los minigenomas en cada pase se cuantificó mediante Q-RT-PCR y se expresó en unidades relativas (gráfico inferior). Se indica en cada caso el minigenoma transfectado en el pase P0. TGEV, células sólo infectadas; M33(-V), células sólo transfectadas con el minigenoma M33; H₂O, control negativo de PCR. (C) Los sobrenadantes del pase P0 de las células BHK-pAPN transfectadas con los minigenomas se pasaron sucesivamente sobre células ST (esquema superior) y la acumulación de los minigenomas en cada pase se cuantificó mediante Q-RT-PCR (gráfico inferior).

minigenoma se sintetizó por transcripción *in vitro* y el porcentaje de células transfectadas con este RNA se determinó mediante citometría de flujo. En las condiciones de transfección optimizadas se consiguieron eficiencias de transfección de RNA próximas al 80% en las células BHK-pAPN previamente infectadas (resultados no mostrados). Posteriormente, se evaluó la replicación en el pase P0 de los minigenomas en las células BHK-pAPN. Para ello se utilizó el minigenoma M33L, con la mutación en la posición 637 corregida por la secuencia consenso viral. Como control negativo se construyó un minigenoma no replicativo mediante la delección de un fragmento de 171 nt de la 3'UTR. Previamente se habían identificado en esta región motivos de estructura secundaria de RNA esenciales para la replicación (Williams y col., 1999). Como control positivo se transfectó el minigenoma M33 sin modificar en la posición 637, que se rescata de forma eficiente y estable por el virus complementador (Izeta y col., 1999) (Fig. 14A). Los distintos minigenomas se transcribieron *in vitro* y se transfectaron sobre células BHK-pAPN, previamente infectadas con el TGEV. Como un control adicional, los minigenomas se transfectaron en células BHK-pAPN no infectadas por el virus complementador. A las 22 hdt (pase P0) se extrajo el RNA y la acumulación de RNA de los distintos minigenomas se cuantificó mediante Q-RT-PCR, utilizando los oligonucleótidos Q-2060-VS y Q-2167-RS (Tabla VIII y Fig. 17), que amplifican la primera región discontinua de secuencia de los minigenomas respecto de la secuencia del genoma viral. No se observaron diferencias significativas en la replicación de los distintos minigenomas (Fig. 14B), debido a la replicación ineficiente en el pase P0, por lo que se decidió analizar el rescate de los minigenomas a lo largo de pases sucesivos. Para ello, los sobrenadantes del pase P0 se pasaron cada 22 h sobre nuevas monocapas de células BHK-pAPN y en cada pase se recogió el RNA y el sobrenadante. Alternativamente, los sobrenadantes de pase P0 procedentes de la transfección de los minigenomas en células BHK-pAPN, se pasaron también sobre células ST, por si las diferencias observadas en la producción de virus entre los dos tipos celulares fueran determinantes para la replicación eficiente de los minigenomas. La amplificación de los minigenomas por el virus complementador se evaluó determinando la acumulación de su RNA, a lo largo de los pases, mediante Q-RT-PCR (Fig. 14B y C). Como control del fondo de RNA procedente de la transfección en pase P0, se cuantificó el RNA del minigenoma M33 a lo largo de los pases en células no infectadas [Fig. 14, M33(-V)]. El fondo inespecífico de hibridación de los oligonucleótidos con la secuencia del genoma viral se determinó a partir de pases de virus de células no transfectadas (Fig. 14,

TGEV). En el caso de los pases sucesivos en células BHK-pAPN (Fig. 14B) se observó que todos los minigenomas, incluido el M33, se perdían a lo largo de los pases, indicando que las células BHK-pAPN no favorecían el rescate eficiente de los minigenomas, a pesar de ser susceptibles a la infección por el TGEV y de su eficiente transfección de RNA (Fig 14B). Sin embargo, en el caso de pasar los sobrenadantes del pase P0 de las células BHK-pAPN transfectadas sobre células ST, sí se observó el rescate del minigenoma M33 (Fig. 14C), indicando que la eficiencia de rescate del minigenoma depende del tipo celular. Sorprendentemente, aún en este caso no se observó rescate del minigenoma M33L con la secuencia corregida en la posición 637, lo que podría ser consecuencia de haber realizado el pase P0 en células BHK-pAPN o de la sustitución del nucleótido original del minigenoma en la posición 637 por la secuencia viral consenso.

El título de virus se determinó en cada pase del rescate de minigenomas en células BHK-pAPN y ST (Fig. 15). En el pase P0 se obtuvieron títulos bajos, probablemente como consecuencia del efecto tóxico de la transfección y la menor confluencia de las células, que aumentaron hasta 1×10^7 ufp/ml a partir del primer pase en las células BHK-pAPN y fueron una unidad logarítmica superiores en los pases en células ST. Esta diferencia podría ser responsable de la distinta capacidad de rescate de los minigenomas en las células ST respecto de las BHK-pAPN por lo que se seleccionaron las células ST para estudios posteriores, con el fin de evaluar el papel del nucleótido 637 en el rescate de minigenomas. Se decidió realizar el rescate de minigenomas desde el pase P0 en células ST, para evitar posibles efectos debidos al tipo celular sobre la amplificación de minigenomas, y para ello la transfección de RNA se optimizó en las células ST hasta conseguir eficiencias próximas al 50%. Sin embargo, cuando se estudió la replicación de los minigenomas en el pase P0, tampoco se obtuvieron diferencias significativas entre los minigenomas que replicaban y los controles no replicativos (datos no mostrados), lo que limitaba los estudios de replicación con los minigenomas a evaluar su rescate a lo largo de los pases. Por otra parte, al igual que ocurría en las células BHK-pAPN, el minigenoma M33L con la secuencia corregida en la posición 637 no se rescató eficientemente, sugiriendo que este nucleótido podría jugar un papel importante en el rescate de los minigenomas del TGEV por el virus complementador.

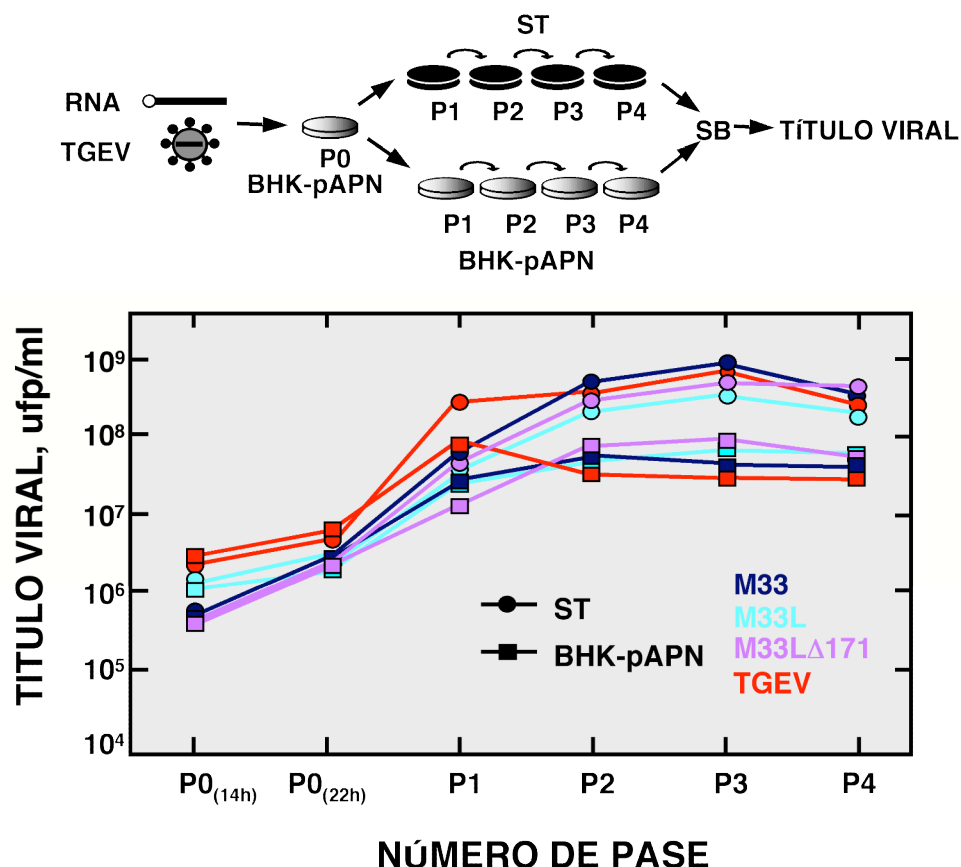


Figura 15. Título de virus durante el rescate de minigenomas. Los minigenomas se transfectaron en células BHK-pAPN infectadas (pase P0) y los sobrenadantes se pasaron en paralelo sobre células ST y BHK-pAPN (esquema superior). En cada pase se recogió una muestra de sobrenadante y se tituló en células ST (gráfico inferior). En el pase P0 se recogieron muestras de sobrenadante a las 14 y 22 hdt, tiempos indicados como subíndices del pase P0 en el eje de abscisas.

1.4. EFECTO DE UNA MUTACIÓN PUNTUAL EN EL GEN DE LA REPLICASA SOBRE EL RESCATE DEL MINIGENOMA M33L

Con el fin de estudiar la relevancia de las mutaciones en el nucleótido 637 sobre la amplificación de los minigenomas, se introdujeron los cuatro posibles cambios de nucleótido en el minigenoma M33L y se analizó su amplificación *in trans* por el virus complementador en células ST. Para ello, se realizaron dos transcripciones *in vitro* de cada uno de los mutantes y se transfectaron por separado en células ST previamente infectadas. Los sobrenadantes de esos cultivos se pasaron cinco veces sobre monocapas de células ST para permitir la amplificación de los minigenomas y la eficiencia de su rescate se determinó mediante Q-RT-PCR, con los oligonucleótidos específicos Q-LBS7-VS y Q-7-RS del minigenoma M33L, que hibridan en la secuencia espaciadora y las secuencias adyacentes, discontinuas respecto del genoma viral (Fig. 16A y tabla

VIII). Para evaluar el grado de transfección de los distintos mutantes, cada duplicado de RNA que se transcribió separadamente, se transfectó por duplicado, se extrajo el RNA total de uno de los duplicados a las 4 hdt, y se utilizó como control para determinar el RNA inicial transfectado. La cantidad de RNA inicial para cada mutante se cuantificó mediante Q-RT-PCR y no se observaron diferencias significativas (Fig. 16B). Adicionalmente, se cuantificó un control de células transfectadas y no infectadas como referencia del fondo de la evaluación [Fig. 16B, M33L(-V)] así como un control de células infectadas pero no transfectadas para comprobar la especificidad de la reacción (Fig. 16B, TGEV). Sólo se rescataron eficientemente los minigenomas M33L con una A o una T en la posición 637, mientras que los minigenomas con una C o una G se perdieron con los pases. Además, en el caso de los minigenomas M33L-637G y M33L-637C se observó un incremento de la acumulación de RNA a partir del tercer pase de virus, lo que parecía indicar una reversión de su fenotipo. Para analizar esta posibilidad, se analizó la secuencia completa de ambos minigenomas en el pase cinco, a partir de fragmentos de RT-PCR amplificados con oligonucleótidos específicos, lo que confirmó la reversión de su secuencia en la posición 637 de G o C a una A en ambos casos. Igualmente se comprobó que las mutaciones introducidas en los minigenomas que se rescataban eficientemente se mantenían a lo largo de los pases (datos no mostrados). Estos resultados indicaban que existía una replicación basal de los minigenomas que rescataban ineficientemente que permitía la selección de revertientes, que la única sustitución del nucleótido 637 era la responsable de los fenotipos observados y que los nucleótidos T o A en la posición 637 eran claramente preferidos para el rescate eficiente del minigenoma M33L por el virus complementador.

Finalmente, para determinar si posibles diferencias en los títulos virales durante los pases seriados eran responsables del rescate diferencial de los minigenomas, los sobrenadantes recogidos en cada pase se titularon en células ST. No se observaron diferencias significativas en los títulos virales que pudieran explicar los fenotipos de replicación de los distintos minigenomas M33L con mutaciones en la posición 637.

1.5. MUTAGÉNESIS Y ANÁLISIS DEL RESCATE DEL RNA DEFECTIVO DI-C

El análisis del rescate de los mutantes del minigenoma M33L mostró que los nucleótidos requeridos en la posición 637 para la amplificación de este minigenoma eran diferentes respecto de los necesarios para el rescate eficiente de virus TGEV a

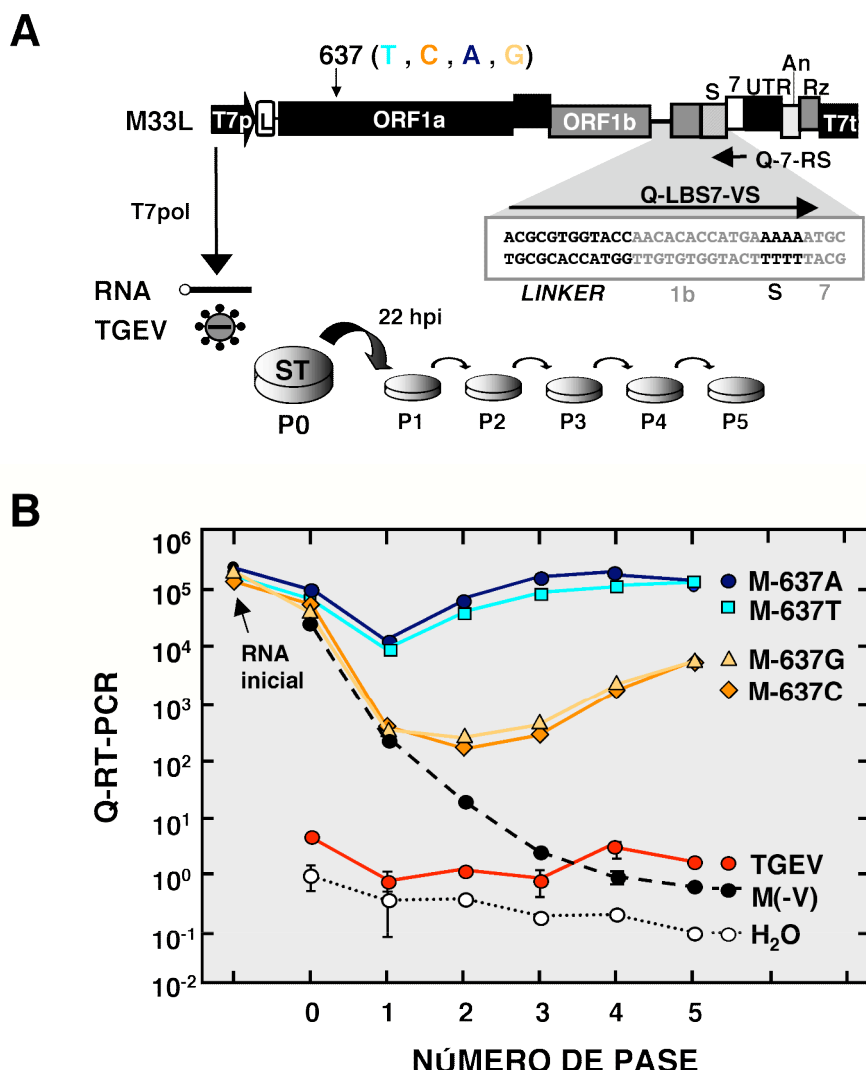


Figura 16. Rescate los mutantes del minigenoma M33L. (A) Esquema experimental. Se representa el cDNA del minigenoma M33L, indicando los genes virales, secuencias flanqueantes y las mutaciones en la posición 637. Los oligonucleótidos utilizados para el análisis por Q-RT-PCR se indican con flechas. La secuencia utilizada para diseñar el oligonucleótido de sentido viral está ampliada debajo del esquema del cDNA del M33L y comprende la secuencia *linker*, específica de este minigenoma y regiones discontinuas del genoma viral. (B) Análisis mediante Q-RT-PCR. La acumulación de RNA de los mutantes se determinó mediante Q-RT-PCR y se expresó en unidades relativas. Los mutantes se indican a la derecha de cada curva con la letra M seguida de la mutación correspondiente. TGEV, células sólo infectadas; M(-V), células sólo transfectadas con el minigenoma M33L-637G; H₂O control negativo. Se indica como RNA inicial la cuantificación de un duplicado de cada transfección de RNA extraído a las 4 hdt. La desviación estándar de replicados de la cuantificación se indica con barras de error.

partir del clon infectivo. Sin embargo, esta característica podría ser específica del minigenoma M33L o bien un requerimiento general para la amplificación de minigenomas derivados del TGEV. Para extender este estudio a un minigenoma de distinto tamaño, el RNA defectivo DI-C de 9.7 kb (Fig. 10) se utilizó para realizar un análisis mutacional similar en la posición 637. Inicialmente se generó el minigenoma

DIC-EGFP para optimizar la transfección de RNAs de gran tamaño, obteniéndose porcentajes de transfección de RNA próximos al 50% en células ST infectadas (datos no mostrados). A continuación se generaron cDNAs que codifican el minigenoma DI-C con cada una de las posibles sustituciones en el nucleótido 637, se transcribieron *in vitro* por duplicado y los RNAs resultantes se transfectaron por separado en células ST infectadas. La amplificación de los mutantes del minigenoma DI-C se determinó mediante Q-RT-PCR a lo largo de los pases, utilizando los oligonucleótidos Q-2060-VS y Q-2167-RS (Fig. 17 y tabla VIII), siguiendo el mismo diseño experimental y controles anteriormente descritos para el análisis de los mutantes del minigenoma M33L.

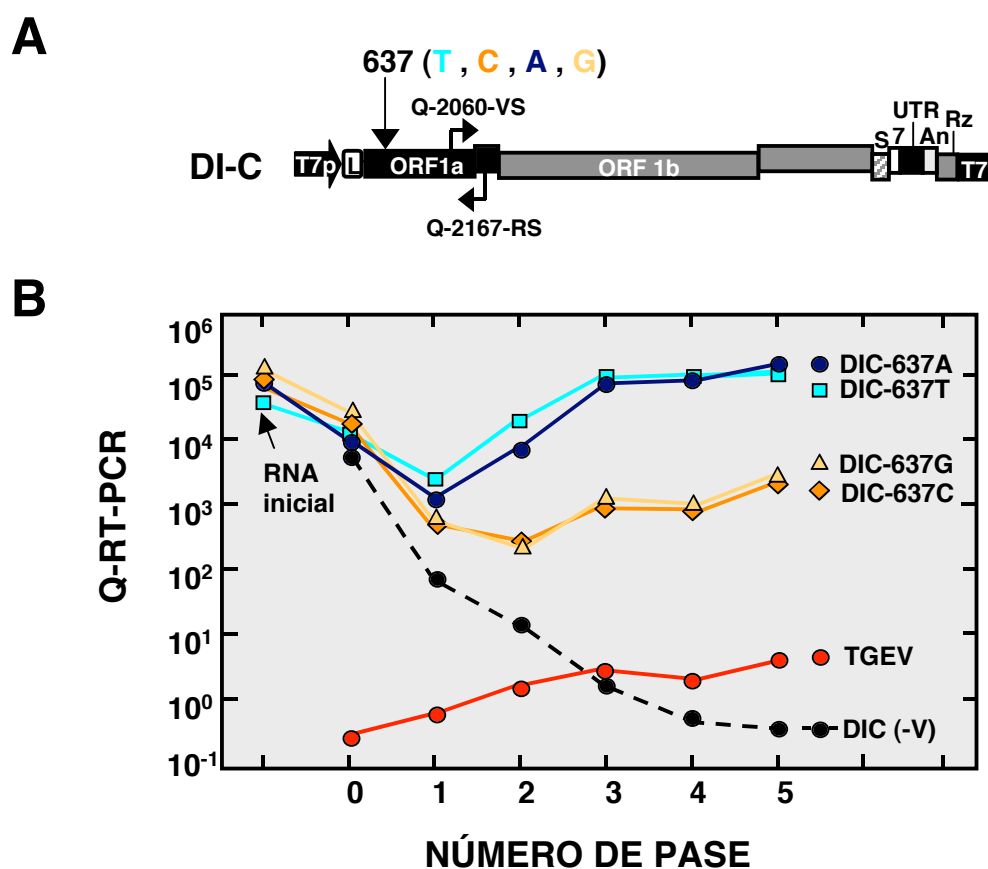


Figura 17. Rescate los mutantes del minigenoma DI-C. (A) Esquema del minigenoma DI-C, mostrando los genes virales y secuencias flanqueantes. Se indican con flechas las mutaciones en el nucleótido 637 y los oligonucleótidos utilizados para el análisis por Q-RT-PCR, diseñados a ambos lados de la primera discontinuidad de secuencia con respecto del genoma viral. T7p y T7t, promotor y terminador de T7, respectivamente; UTR, región 3' no traducida; An, poliA; Rz, secuencia de la ribozima del HDV. (B) Q-RT-PCR. La acumulación del RNA de los mutantes del RNA DI-C a lo largo de los pases se cuantificó mediante Q-RT-PCR y se representa en unidades relativas. Como controles se cuantificaron muestras de células sólo infectadas (TGEV) y células sólo transfectadas con el RNA DI-C-637T [DIC (-V)]. H₂O, control negativo de la PCR. Se indica como RNA inicial la cuantificación de un duplicado de cada transfección de RNA extraído a las 4 hdt.

Los mutantes del minigenoma DI-C con una T o una A en la posición 637 se rescataron eficientemente a lo largo de los pases, al contrario que aquellos que contenían una C o una G en la misma posición, mostrando un comportamiento similar al de los mutantes correspondientes del minigenoma M33L. Para analizar posibles diferencias debidas a la infección, se determinó el título de virus en los sobrenadantes de todos los pases, sin que se observaran diferencias que pudieran explicar los fenotipos observados durante el rescate de los RNAs defectivos (resultados no mostrados). Los resultados demostraban los mismos requerimientos de secuencia en la posición 637 para el rescate de los minigenomas M33L y DI-C, a pesar de su distinto tamaño.

1.6. EFECTO DE UNA MUTACIÓN PUNTUAL EN EL GEN DE LA REPLICASA SOBRE PREDICCIONES DE ESTRUCTURA SECUNDARIA DEL RNA

Los resultados anteriores indicaban que, tanto la replicación en *cis* del genoma del TGEV como el rescate en *trans* de minigenomas derivados, estaba afectada por las mutaciones puntuales del nucleótido 637, y que en esta posición se seleccionaban distintos nucleótidos para el rescate eficiente del virus o de minigenomas. Las mutaciones en la posición 637 producen un cambio de aminoácido en la posición 108 de las poliproteínas pp1a y pp1ab de la replicasa (Fig. 12A). El aminoácido 108 se encuentra en el sitio potencial de procesamiento N-terminal de las poliproteínas, entre los polipéptidos p9/p87 (Fig. 7), según predicciones basadas en la comparación de secuencia de la replicasa del TGEV con la del HCoV-229E (Herold y col., 1998). El análisis mutacional realizado con el cDNA de longitud completa del TGEV mostró una correlación entre los cambios más drásticos de aminoácido y aquellos virus más afectados en replicación, lo que sugería que el efecto de las mutaciones en la posición 637 era a nivel de proteína. Sin embargo, no se podía descartar la posibilidad de que cambios en la estructura secundaria del motivo de RNA que contiene el nucleótido 637 pudieran ser responsables de los distintos fenotipos observados en los virus y minigenomas. Con el fin de analizar esa posibilidad, se analizó la estructura secundaria del RNA del TGEV en el entorno del nucleótido 637 utilizando el programa *M-Fold* v.3.1 (Mathews y col., 1999; Zuker, 2003). Se realizaron predicciones estructurales a partir de fragmentos del extremo 5' del genoma del TGEV de 500 nt, centrados en la posición 637, con cada uno de los posibles cambios en dicha posición. Se obtuvieron dos modelos de plegamiento, uno común para las construcciones que contenían una A o

Resultados

una G en la posición 637 (Fig. 18A, estructura A) y otro común a las construcciones que contenían una C o una T en la misma posición (Fig. 18A, estructura B). Un análisis posterior más detallado, considerando secuencias de diferente longitud en torno a la posición 637, así como el análisis de las secuencias complementarias de polaridad negativa, fue consistente con estas predicciones estructurales (resultados no mostrados). En base a estas predicciones de estructura, se analizó la posible correlación entre las diferentes mutaciones en la posición 637, la replicación de los virus o minigenomas analizados y la estructura secundaria del RNA en cada caso (Fig. 18B). No se encontró ninguna correlación entre los posibles cambios de estructura del RNA y los fenotipos asociados a las distintas mutaciones, lo que apoyaba la hipótesis alternativa de que las mutaciones en la posición 637 estaban afectando a nivel de proteína.

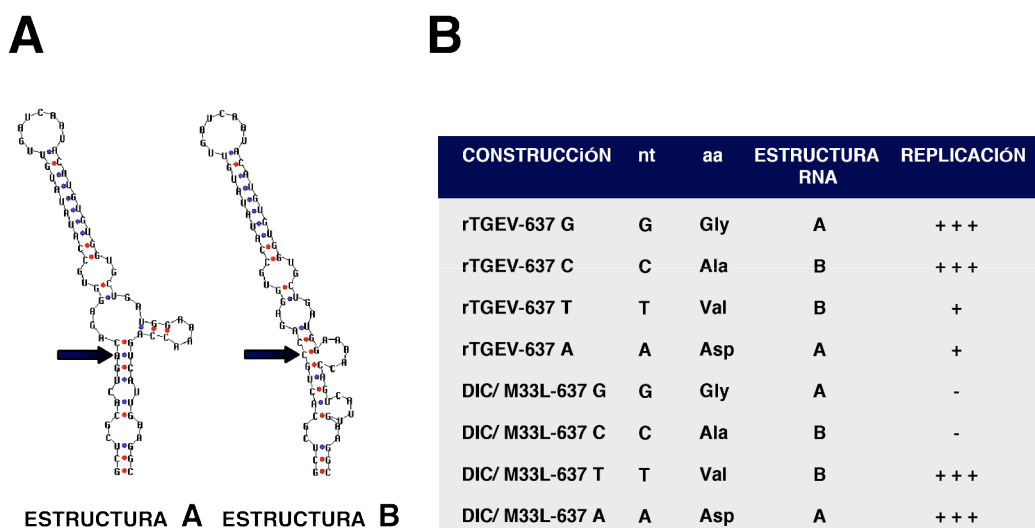


Figura 18. Efecto de las mutaciones en la posición 637 sobre predicciones de estructura secundaria del RNA del TGEV. (A) Estructura secundaria del RNA del TGEV con mutaciones en la posición 637. Se realizaron las predicciones de estructura secundaria de RNA de un fragmento de 500 nt del genoma del TGEV utilizando el programa *M-fold v.3.1*. Se muestra el motivo de estructura que contiene la posición 637, que presentaba un plegamiento común con las sustituciones del nucleótido 637A y 637G (estructura A) y otro con las mutaciones 637T y 637C (estructura B). (B) Correlación entre las predicciones de estructura secundaria de RNA con las mutaciones en la posición 637, el cambio de aminoácido asociado a cada mutación, y los fenotipos observados en el virus y minigenomas con las mutaciones correspondientes.

1.7. EFECTO DE UNA MUTACIÓN PUNTUAL SOBRE EL PROCESAMIENTO *IN VITRO* NSP1/NSP2 DE LAS POLIPROTEÍNAS DE LA REPLICASA POR LA PROTEINASA VIRAL PLP1

En conjunto, los resultados anteriores indicaban que el efecto de las mutaciones en la posición 637 del genoma del TGEV sobre la replicación del virus podía deberse a un defecto del procesamiento N-terminal de las poliproteínas de la replicasa. El sitio de corte N-terminal de la proteasa PLP1 ocupa una posición variable en el extremo de las poliproteínas de diferentes coronavirus (Fig. 7) (Herold y col., 1998; Hughes y col., 1995). La proteasa PLP1 del coronavirus humano HCoV-229E, estrechamente relacionado con el coronavirus TGEV, produce un corte entre los aminoácidos Gly111 y Asn112 que libera un polipéptido N-terminal de 9 kDa (Fig. 7). Hasta la fecha no se han publicado datos acerca del procesamiento N-terminal de las poliproteínas del TGEV, aunque se han propuesto dos sitios de corte potenciales, por análisis comparativo de secuencia con el coronavirus HCoV-229E. De acuerdo con estas predicciones, el nucleótido de la posición 637 ocupa la segunda posición del codon correspondiente al aminoácido 108, que correspondería a la posición P1' respecto del sitio potencial de corte por la proteasa PLP-1 Thr107/Gly108 o bien la posición P3 respecto del sitio de procesamiento alternativo Gly110/Ala111 (Fig. 19), según la nomenclatura establecida por Schechter y Berger para los sitios de corte proteolítico (NH₂-P3-P2-P1↓P1'-P2'-P3'-COOH). En ambos casos, un cambio drástico de aminoácido en dicha posición podría afectar al procesamiento mediado por la enzima, según lo descrito a partir del análisis de secuencia y mutagénesis de tres sitios de procesamiento de la proteasa PLP1 en otros coronavirus (Bonilla y col., 1997; Dong y Baker, 1994; Hughes y col., 1995). Para analizar esta hipótesis, se realizó un ensayo de procesamiento *in vitro*. Para ello se sintetizaron como sustrato dos fragmentos, comprendiendo los primeros 610 aminoácidos del extremo N-terminal de las poliproteínas, con una glicina o un aspártico en el aminoácido 108, que es el cambio más drástico de aminoácido producido por la mutación de la posición 637 del genoma (637G por A). Se utilizaron sustratos de 610 aminoácidos porque es el fragmento N-terminal de la replicasa común en el genoma y los minigenomas M33L y DI-C, los cuales codifican proteínas de fusión formadas por regiones discontinuas de secuencia con respecto al genoma viral (Fig. 41). La secuencia de la proteasa PLP1 del TGEV se clonó incluyendo el dominio similar a la papaína y secuencias adyacentes necesarias para su actividad proteasa, por homología con ensayos de actividad en *trans* realizados con mutantes de la proteína equivalente del HCoV-

Resultados

229E (Herold y col., 1998) (Fig. 19A). Las construcciones que contienen las secuencias de la enzima y los sustratos se clonaron bajo el promotor del fago T7 y se utilizaron por separado como molde de reacciones de transcripción-traducción, en las que los sustratos se marcaron radiativamente con ^{35}S -Metionina. Las reacciones de traducción se pararon añadiendo cicloheximida y seguidamente se realizó el ensayo de procesamiento *in trans*, combinando las mezclas de reacción de la enzima y los sustratos en distintas

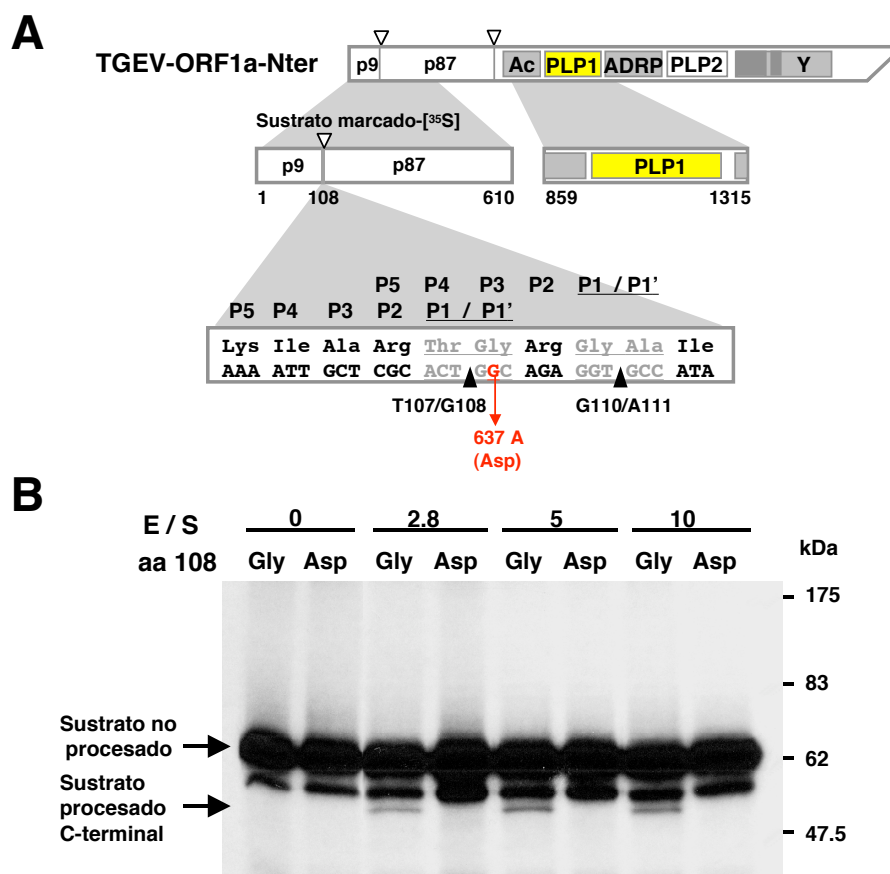


Figura 19. Efecto de mutaciones puntuales en el nt 637 sobre el procesamiento in N-terminal de las poliproteínas de la replicasa. (A) Esquema del extremo N-terminal de las poliproteínas del TGEV (barra superior) indicando los fragmentos utilizados para el procesamiento *in vitro* (barras inferiores). Los números debajo de las barras indican posiciones de aminoácido. Los sitios de procesamiento se indican con puntas de flecha invertidas. Se muestra ampliada la secuencia del sitio de corte N-terminal indicando los dos posibles sitios de procesamiento, por homología de secuencia con el HCoV-229E. Los aminoácidos flanqueantes a los sitios de procesamiento se denominan según la nomenclatura NH_2 -P3-P2-P1↓P1'-P2'-P3'-COOH (Schechter y Berger, 1967). El nucleótido de la posición 637 está señalado con una flecha y se indica la mutación G por A y el cambio de aminoácido que produce. ADPR, actividad adenosina difosfato ribosa 1'' fosfatasa; Ac, dominio ácido; Y, dominio que contiene una región transmembrana y un dominio potencial de unión a metales. (B) Ensayo de procesamiento *in trans*. Dos sustratos del extremo N-terminal de 610 aminoácidos conteniendo una Gly o Asp en la posición 108 se sintetizaron por transcripción-traducción *in vitro* en presencia de ^{35}S -Metionina y se incubaron en presencia de la proteasa PLP1, sintetizada en ausencia de marca. Las incubaciones se realizaron a 30°C durante 16 h, con las relaciones enzima-sustrato (E/S) indicadas. Los productos se resolvieron mediante 10% PAGE-SDS y se revelaron por fluorografía. Se indican con flechas a la izquierda las formas procesada y no procesada del sustrato; los marcadores de masa molecular, en kDa, se muestran a la derecha.

proporciones, incubándose a 30°C durante 14 h (Fig. 19B). En ausencia de la enzima, se detectó el sustrato marcado con la masa molecular esperada de 67 kDa junto con una banda de menor masa molecular, que podría ser un producto de terminación prematuro de la transcripción o de la traducción. Sin embargo, en presencia de la proteasa PLP1, se detectó una forma procesada del sustrato en el caso de la construcción que presentaba un residuo de glicina en la posición del aminoácido 108, pero no con el sustrato que contenía un aspártico en la misma posición. El tamaño del producto de procesamiento coincidió con el esperado para el fragmento de procesamiento C-terminal del sustrato generado después de la liberación del polipéptido N-terminal de 9 kDa, y no fue debido a un evento de proteólisis o degradación espontánea de este sustrato particular, puesto que la forma procesada no se detectó en reacciones de incubación de los sustratos en paralelo en ausencia de la enzima. Sin embargo, no se detectó el producto de procesamiento N-terminal de 9 kDa debido probablemente a que solo contiene dos residuos de metionina y a la reducida eficiencia de corte de la enzima PLP1. La eficiencia de procesamiento observada fue similar en las diferentes proporciones enzima:sustrato ensayadas. Se trató de optimizar la eficiencia de procesamiento variando la temperatura, el tiempo de incubación y añadiendo ZnCl_2 durante la traducción de la enzima para favorecer el plegamiento correcto de la proteasa, aunque no se consiguió una mejora significativa del procesamiento. A pesar de la limitada eficiencia del procesamiento *in vitro*, los resultados indicaban claramente que la mutación del nucleótido 637 que producía un cambio drástico de aminoácido afectaba al corte de las poliproteínas del TGEV por la proteasa PLP1.

Los distintos requerimientos de secuencia en la posición 637 para la replicación del genoma viral y la amplificación de minigenomas permitían establecer una correlación entre los cambios de nucleótido que predeciblemente permitían el procesamiento N-terminal de las poliproteínas de la replicasa y aquellos virus que eran rescatados eficientemente, indicando que este procesamiento era esencial para el ciclo infectivo del virus. Sin embargo los minigenomas seleccionaban mutaciones que probablemente afectaban *in vivo* a este sitio de corte. Estas observaciones sugerían la existencia de una estrategia diferente, desarrollada por los minigenomas, para su amplificación en *trans* dependiente del virus complementador. Estas posibles diferencias de la replicación en *cis* y en *trans* podían limitar el uso de los minigenomas para estudios básicos de replicación cuyas conclusiones quisieran ser extrapoladas al genoma viral. Además, en los estudios de rescate de los minigenomas no se había podido analizar la replicación en

Resultados

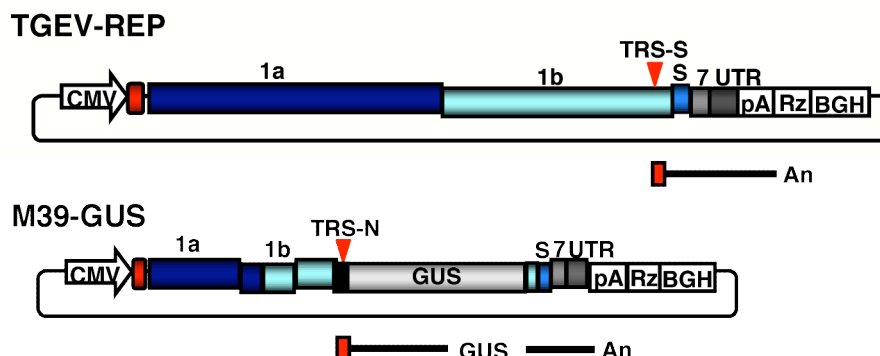
pase P0 de forma independiente de la encapsidación. Con el objetivo de generar un sistema con el que poder abordar estudios básicos de replicación *in vivo*, se decidió generar un replicón derivado del TGEV. Esta aproximación permitiría evaluar la replicación de forma independiente al proceso de encapsidación.

2. GENERACIÓN Y CARACTERIZACIÓN DE UNA COLECCIÓN DE REPLICONES DERIVADOS DEL GENOMA DEL TGEV

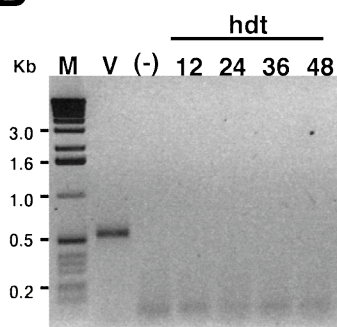
2.1. GENERACIÓN DE UN REPLICÓN DEL TGEV QUE CONTIENE LAS SECUENCIAS MÍNIMAS TEÓRICAS PARA SU REPLICACIÓN AUTÓNOMA

En primer lugar, se construyó un replicón derivado del TGEV basado en los datos publicados para distintos virus RNA de cadena positiva, entre los que se incluye el arterivirus equino (EAV), para los que únicamente el gen de la replicasa es suficiente para la replicación autónoma del RNA viral (Frolov y col., 1996; Khromykh y Westaway, 1997; Lohmann y col., 1999; Molenkamp y col., 2000). Este primer replicón (TGEV-REP), que contenía los extremos 5' y 3' del genoma y el gen de la replicasa, se clonó en un BAC bajo el control del promotor de CMV y se flanqueó en su extremo 3' con las secuencias necesarias para la correcta terminación y procesamiento del RNA dentro de la célula (Fig. 20A). Debido al solapamiento del gen de la replicasa y el gen S, el replicón contenía la TRS y los tres primeros codones del gen de la proteína S, que daría lugar por transcripción a un RNA subgenómico formado por fragmentos discontinuos de los genes S y 7 junto con el extremo 3' del genoma. Para evaluar la actividad en *cis* del replicón, éste se transfectó por cuatriplicado en células ST con el reactivo *Lipofectin* (Invitrogen) y el RNA total se recogió a los tiempos indicados (Fig. 20B). La presencia del transcrito generado a partir de la TRS del gen S se analizó mediante RT-PCR, sin que se detectara actividad significativa del replicón. En paralelo se analizó la actividad en *trans* del replicón, utilizando un sistema muy sensible basado en la expresión del gen trazador de la β -glucuronidasa (GUS) a partir del minigenoma M39 (Alonso y col., 2002). El minigenoma M39 está derivado del RNA DI-C (Izeta y col., 1999) y se replica eficientemente por el virus complementador. El cDNA de este minigenoma estaba clonado bajo el promotor de CMV y posteriormente se insertó el gen GUS bajo el control de la TRS del gen N, generando el minigenoma M39-GUS (Fig. 20A). Para analizar la actividad en *trans* del replicón, el minigenoma M39-GUS se cotransfectó junto con el replicón TGEV-REP en células ST utilizando el reactivo *Lipofectin* (Invitrogen). En paralelo se utilizaron, como control positivo, células transfectadas con el minigenoma e infectadas con el TGEV como virus complementador.

A



B



C

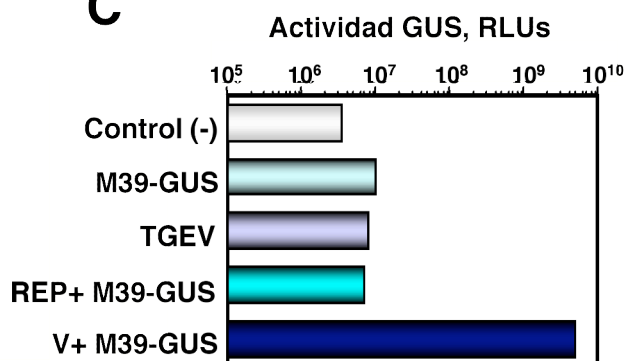


Figura 20. Ensayo de la actividad del replicón mínimo teórico TGEV-REP. (A) Esquema de las construcciones TGEV-REP y M39-GUS, indicando en color los genes virales y en blanco las secuencias flanqueantes, para la generación correcta del RNA en las células. CMV, promotor de citomegalovirus; pA, poliA; Rz, secuencia de la ribozima del HDV; BGH, secuencia de terminación y poliadenilación del gen de la hormona BGH. Las TRSs se indican encima de los genes con flechas invertidas y los RNAs subgenómicos se representan por líneas gruesas debajo de las construcciones, destacando la secuencia líder en el extremo 5' (caja roja) y el poliA (An) en el extremo 3'. (B) Análisis de la actividad *in cis* del replicón mínimo teórico TGEV-REP. El replicón TGEV-REP se transfectó en células ST con el reactivo *Lipofectin*, se extrajo el RNA a los tiempos indicados y se analizó el mensajero generado por la TRS del gen S mediante RT-PCR. V, infección con el virus TGEV, como control positivo; M, células no infectadas, como control negativo. (C) Análisis de la actividad *in trans* del replicón mínimo TGEV-REP. La actividad *in trans* se evaluó cuantificando la expresión del gen *gus* desde el minigenoma M39 en presencia del replicón. La actividad GUS se muestra en unidades luminométricas relativas (RLUs), comparando el fondo de células sin tratar [control (-)], células transfectadas con el minigenoma (M39-GUS), células infectadas (TGEV), células cotransfectadas con el replicón y el minigenoma (REP+M39-GUS) y, como control positivo, células infectadas y transfectadas con el minigenoma (V+M39-GUS).

La amplificación del minigenoma M39-GUS se determinó mediante la expresión de la actividad de la β -glucuronidasa o actividad GUS. No se detectó un incremento significativo de actividad GUS en presencia del replicón, en comparación con los niveles de fondo de los controles negativos, mientras que sí se observó incremento de actividad en presencia del virus complementador (Fig. 20C).

La ausencia de actividad del replicón, tanto en *cis* como en *trans*, podía ser debida a la baja eficiencia de transfección en las células ST de los cDNAs de gran tamaño, o bien a la necesidad de otras secuencias o genes adicionales para la funcionalidad de la replicasa. Para determinar la relevancia de estas variables, se decidió optimizar la transfección de los DNAs de gran tamaño en células de mamífero, así como analizar posibles requerimientos de secuencia adicionales para la actividad del replicón.

2.2. OPTIMIZACIÓN DE LA TRANSFECCIÓN DE DNAs DE ALTA MASA MOLECULAR EN CÉLULAS EUCARIOTAS

La baja eficiencia de transfección de los BAC en células ST (menor del 0,1%) con el reactivo *Lipofectin* (Invitrogen) podía ser un factor limitante para obtener un replicón funcional del TGEV. Aunque en estas condiciones el rescate de virus a partir de la transfección del clon infectivo es eficiente debido al proceso de amplificación viral por infección, la baja eficiencia de transfección podría ser un parámetro crítico para detectar la actividad de un replicón que no genera partículas infectivas. Para optimizar la eficiencia de transfección de BACs con un tamaño próximo a 40 kb, se generó un cDNA no replicativo a partir del clon infectivo del TGEV, en el que se insertó un módulo de expresión conteniendo la secuencia codificante de la EGFP bajo el control del IRES del virus de la encefalomiocarditis, interrumpiendo la fase de lectura de la replicasa. Adicionalmente, se introdujo la mutación puntual en el extremo 5' del genoma que afectaba drásticamente a la replicación viral en la posición 637(A), y el fragmento *ClaI-ClaI* (nt 4417 a 9615) de la replicasa se insertó en la orientación

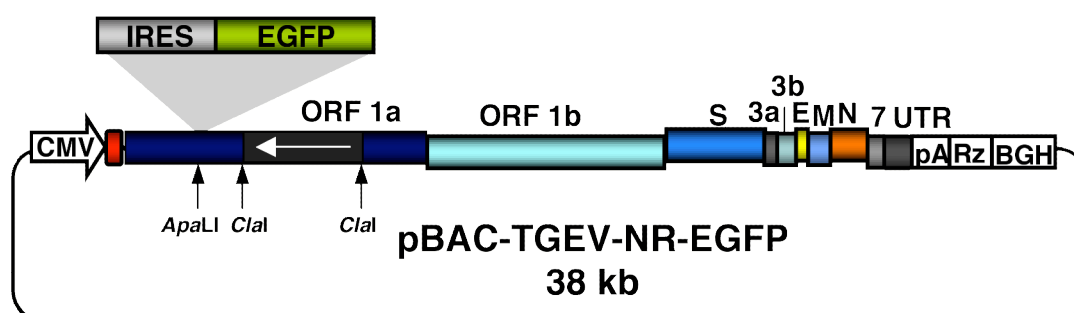


Figura 21. Esquema de la construcción pBAC-TGEV-NR-EGFP utilizada para optimizar la transfección de DNA de alta masa molecular. Esquema del cDNA del TGEV, en el que se clonó el gen de la EGFP precedido por una secuencia IRES en el sitio de restricción *ApaLI*, del gen de la replicasa (ORF1a). Para generar una construcción no replicativa, el fragmento *ClaI-ClaI* de la ORF1a se insertó en la orientación inversa. Se indican los genes virales y secuencias flanqueantes. CMV, promotor de citomegalovirus; pA, poliA; Rz, secuencia de la ribozima del HDV; BGH, secuencia de terminación y poliadenilación del gen de la hormona BGH.

Resultados

inversa. De esta forma se generó el plásmido pBAC-TGEV-NR-EGFP, de 38 kb (Fig. 21), que permite evaluar la eficiencia de transfección mediante citometría de flujo, analizando el porcentaje de células que expresan la proteína EGFP. Para optimizar la transfección se utilizaron células de hámster BHK-pAPN y humanas 293T, puesto que

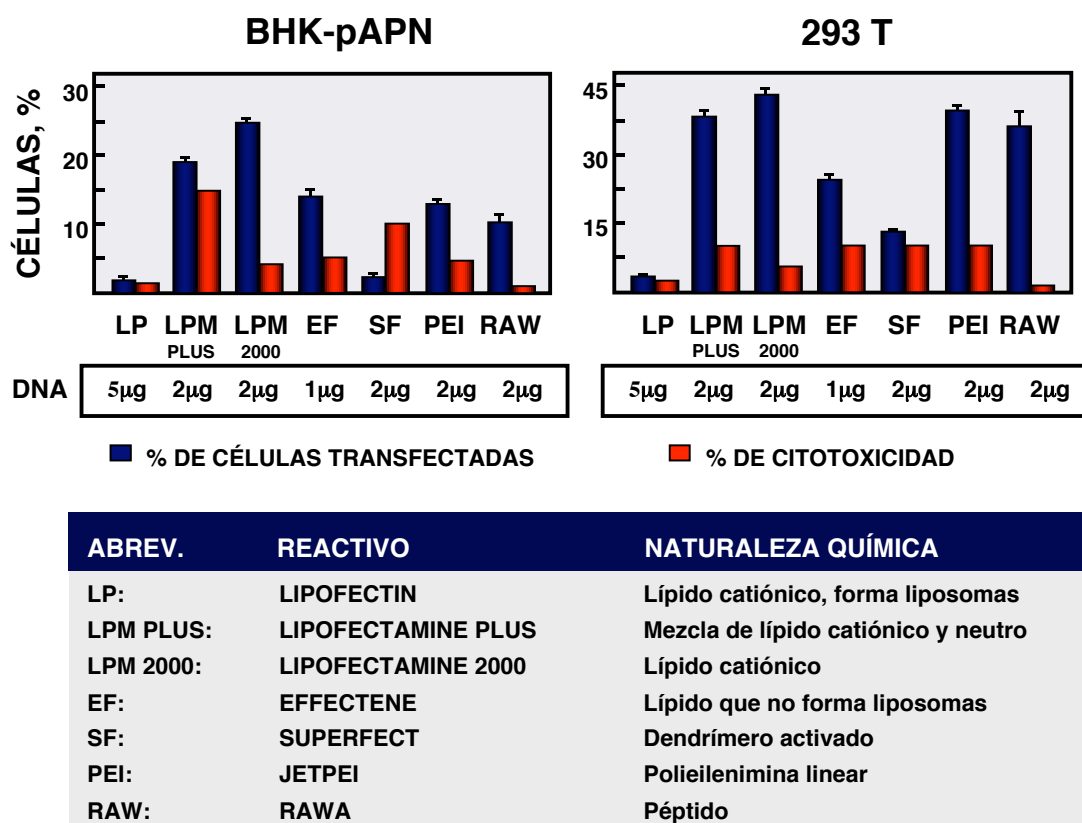


Figura 22. Comparación de reactivos para la transfección de BACs en células de mamífero. La eficiencia de transfección de la construcción pBAC-TGEV-NR-EGFP se evaluó en paralelo en células BHK-pAPN y 293T, utilizando siete agentes de transfección comerciales. El porcentaje de células transfectadas se determinó mediante citometría de flujo, cuantificando el porcentaje de células fluorescentes a las 20-24 hpt. El % de citotoxicidad indica el porcentaje aproximado de células muertas. En todos los casos la transfección se llevó a cabo en placas de 35 mm de diámetro siguiendo las condiciones especificadas por el fabricante. La cantidad de DNA transfectado en cada caso se indica en los recuadros debajo del eje de abscisas. En el cuadro inferior se indican las abreviaturas del nombre comercial de los reactivos de transfección y su naturaleza química.

ambas líneas celulares soportan la replicación del virus cuando este es introducido en forma de cDNA y permiten eficiencias de transfección 10 y 20 veces superiores a las células ST, respectivamente. Se evaluaron siete reactivos de transfección de distinta naturaleza química, en ambos tipos celulares, siguiendo en cada caso las condiciones de transfección estándar indicadas por el fabricante, y se evaluó el porcentaje de células transfectadas, así como el efecto citotóxico estimado de la transfección, expresado como

el porcentaje de células muertas respecto de un control sin transfectar (Fig. 22). Un conjunto de los agentes de transfección probados está formado por lípidos catiónicos que interaccionan con los grupos fosfato de los ácidos nucleicos y forman liposomas (Bertling y col., 1991; Felgner y col., 1987), constituido por los reactivos *Lipofectin* (Invitrogen), *Lipofectamine Plus* (Invitrogen) y *Lipofectamin 2000* (Invitrogen). El reactivo *Effectene* (QIAGEN) también es de naturaleza lipídica pero no forma liposomas. Entre los reactivos no lipídicos se utilizó el *JetPEI* (Polytransfection) que consiste en una polietilenimina lineal que compacta el DNA en partículas cargadas positivamente, que interaccionan a su vez con los proteoglicanos de la superficie celular y entran en las células por endocitosis (Boussif y col., 1995; Mislick y Baldeschwieler, 1996). Otros agentes no lipídicos que se probaron fueron el reactivo *SuperFect* (QIAGEN), que es un dendrímero activado de estructura ramificada (Tang y col., 1996) y el péptido *RAWA* (Medplant Genetics) de naturaleza catiónica y con estructura de alfa hélice anfipática, lo que le permite interaccionar con el DNA cargado negativamente, compactarlo e introducirlo en las células (Fominaya y col., 2000).

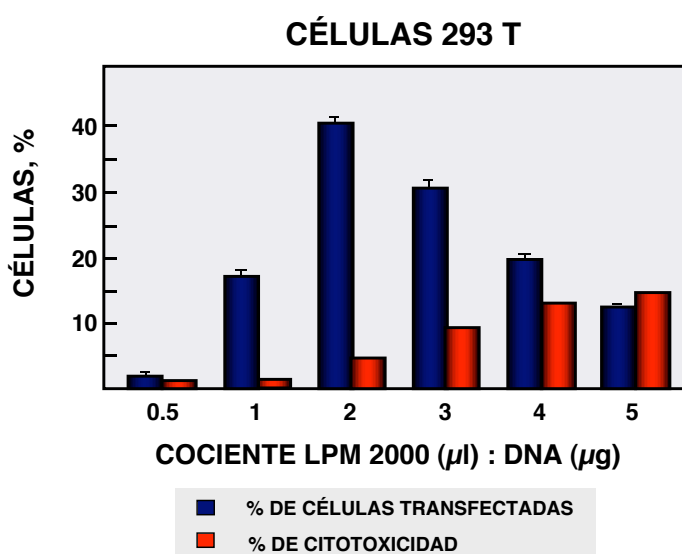


Figura 23. Optimización de la transfección de BAC en células 293T. Se evaluaron diferentes relaciones entre el reactivo seleccionado, *Lipofectamine 2000* (LPM 2000), y la cantidad de DNA del plásmido pBAC-TGEV-NR-EGFP. A las 20-24 hdt se evaluó el porcentaje de transfección mediante citometría de flujo y se estimó el porcentaje de células muertas.

La mejor eficiencia de transfección, junto con niveles de citotoxicidad reducidos, se consiguió utilizando el reactivo *Lipofectamin 2000* (Fig. 22), por lo que se seleccionó este reactivo para una optimización posterior. Uno de los parámetros clave a optimizar con los reactivos de lipofección para conseguir una mejor eficiencia de transfección con un menor daño celular es la proporción lípido:ácido nucleico. Para ajustar esta relación se utilizaron las células 293T que se transfectaron con proporciones crecientes de

Lipofectamine 2000 respecto de plásmido pBAC-TGEV-NR-EGFP y se determinó el porcentaje de transfección y de citotoxicidad (Fig. 23). La máxima eficiencia de transfección con un bajo nivel de citotoxicidad se obtuvo con una relación de reactivo de lipofección (μl) /DNA (μg) igual a dos (Fig. 23), alcanzándose niveles de transfección del 42%. Utilizando estas condiciones se consiguieron niveles de transfección del 20% en el caso de las células BHK-pAPN (datos no mostrados).

2.3. EFECTO DEL AUMENTO DE LA EFICIENCIA DE TRANSFECCIÓN SOBRE DEL RESCATE DE VIRUS A PARTIR DEL CLON INFECTIVO DEL TGEV

Una vez optimizada la eficiencia de transfección de los BAC en células de mamífero, se analizó el efecto de esta variable sobre la recuperación de virus infectivo a partir del clon infectivo del TGEV. Para ello se transfectó el plásmido pBAC-TGEV^{FL} con las condiciones de transfección iniciales y las optimizadas, y se analizó el rescate de virus a los tiempos indicados (Fig. 24). La transfección de las células ST con el reactivo *Lipofectin* (condiciones de transfección iniciales) dio lugar a la recuperación de sólo 10^2 ufp/ml a las 72 hdt. Para incrementar la eficiencia de transfección y la producción de virus, se utilizaron células BHK-pAPN con este mismo reactivo, que permiten eficiencias de transfección unas 10 veces superiores respecto de las ST, y seis horas después de la transfección las células se levantaron con tripsina y se sembraron sobre una monocapa de células ST, donde el virus replica más eficientemente, para incrementar la producción viral. Esta estrategia produjo un incremento de cuatro unidades logarítmicas en la producción de virus infectivo determinado a las 72 hpt (Fig. 24). Esta misma estrategia se utilizó con el reactivo *Lipofectamin 2000* y las condiciones optimizadas de transfección, tanto en células BHK-pAPN como en 293T, lo que dio lugar a una producción de virus cercana a 10^8 ufp/ml a las 72 hpt con ambos tipos celulares.

Como resultado de esta estrategia basada en transfectar células BHK o 293T con *Lipofectamine 2000* y posteriormente añadir las células transfectadas sobre monocapas de células ST, se consiguió un aumento del rescate de virus infectivo en pase P0 de cinco a seis unidades logarítmicas.

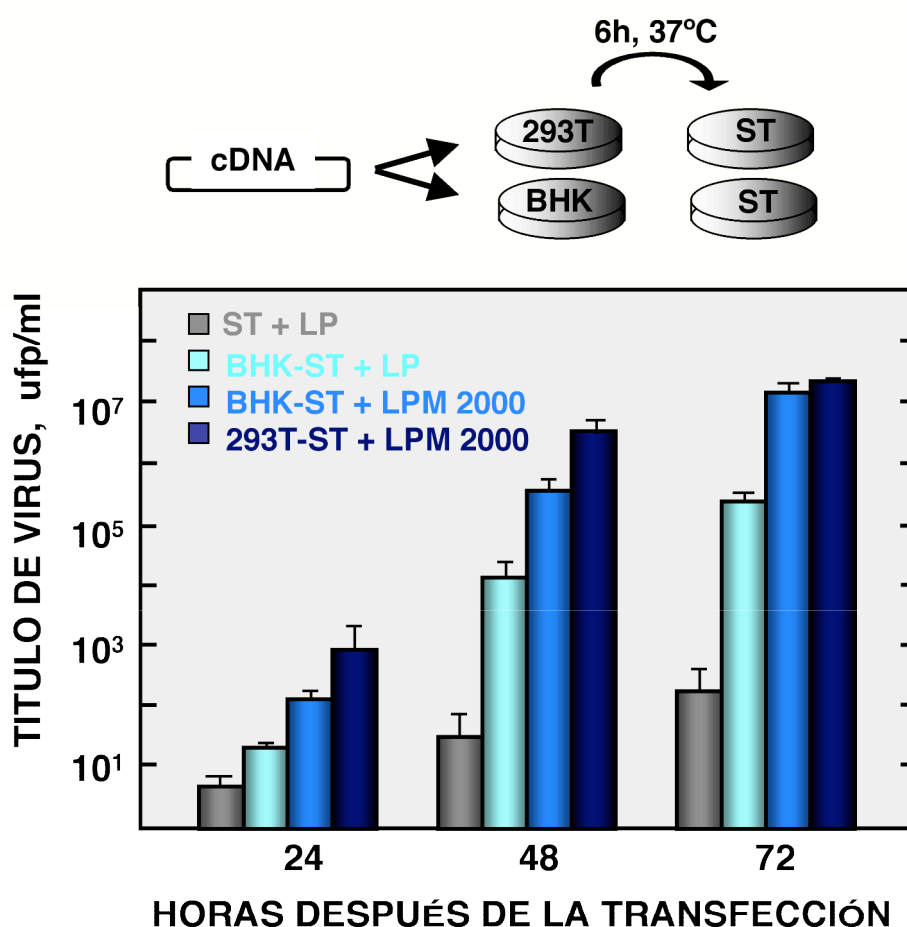


Figura 24. Mejora de la eficiencia de rescate de virus infeccioso como consecuencia de la optimización de la transfección de los BACs. En el esquema superior se muestra la estrategia de utilizar dos tipos celulares para mejorar el rescate de virus, en un primer paso células 293T o BHK muy permisivas para la transfección, y en un segundo paso células ST sobre las que se añaden las células transfectadas para amplificar la producción de virus. En la gráfica se muestra el título de virus obtenido a partir de células ST, o bien con la estrategia BHK-ST o 293T-ST. Los sobrenadantes de estos cultivos se recogieron a los tiempos indicados después de la transfección y se titularon en células ST. Se compararon en paralelo los títulos obtenidos con las condiciones iniciales de transfección con *Lipofectin* (LP) o con las condiciones optimizadas de transfección con *Lipofectamine 2000* (LPM 2000).

2.4. GENERACIÓN DE REPLICONES FUNCIONALES DERIVADOS DEL TGEV

Una vez optimizada la transfección de los BAC se generó una colección de replicones derivados del TGEV para estudiar el papel de genes distintos de la replicasa, sobre la funcionalidad del replicón. Estudios anteriores en el laboratorio habían demostrado que los genes 3a, 3b y 7 no son esenciales para la replicación del virus en cultivos celulares y que los genes S y E, aunque son esenciales para la morfogénesis del virus, no afectan a la replicación del genoma (Ortego y col., 2002; 2003; Sola y col., 2003). Para la construcción de los replicones del TGEV se utilizó un clon infeccioso

modificado del TGEV (TGEV-RS) en el cual los genes se habían separado mediante la duplicación de las secuencias solapantes y se habían introducido sitios únicos de restricción en el extremo 5' de cada gen para facilitar la manipulación del genoma (Ortego y col., 2003). Partiendo de este clon infectivo modificado se generó una colección de tres replicones del TGEV mediante delección sucesiva de genes (Fig. 25). El replicón REP 1 contenía los genes M, N y 7, además del gen de la replicasa y las

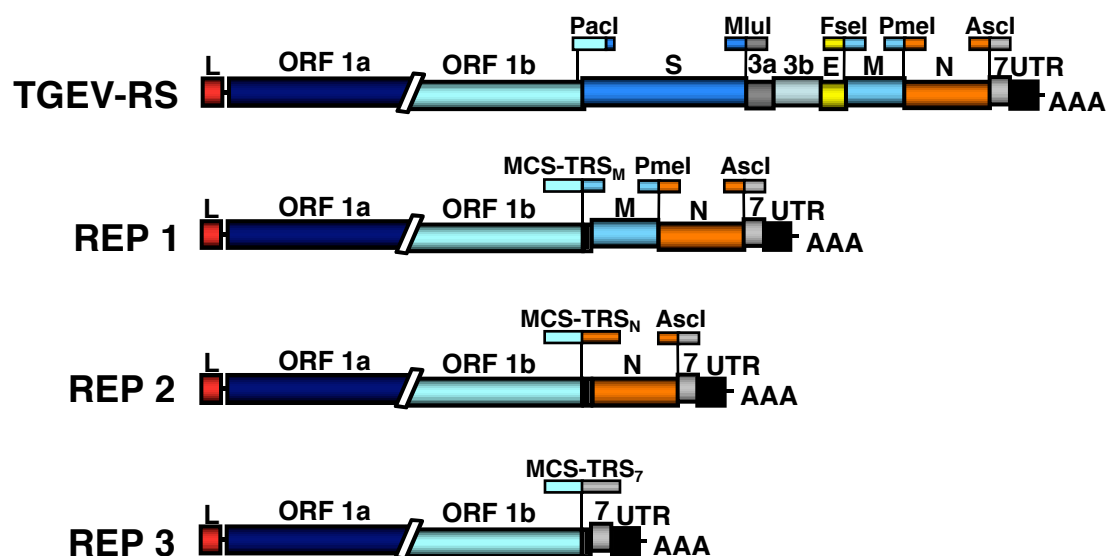


Figura 25. Replicones derivados del TGEV mediante delección sucesiva de genes. A partir de un cDNA infectivo del TGEV con los genes separados por dianas únicas de restricción (TGEV-RS, barra superior), se generó una colección de replicones que contenían, además del gen de la replicasa, los genes M, N y 7 (REP 1), los genes N y 7 (REP 2) o únicamente el gen 7 (REP 3). En los tres replicones se introdujo un sitio múltiple de clonaje (MCS) para permitir la inserción posterior de genes heterólogos. En cada construcción se indican los genes virales y las dianas únicas de restricción. L, secuencia líder; UTR, región 3' no traducida; AAA, poliA.

secuencias de los extremos 5' y 3' del genoma. El replicón REP 2 se generó mediante la delección posterior del gen M, y el replicón REP 3 mediante la delección de los genes M y N. El gen 7 se mantuvo en los tres replicones ya que podía contener en su secuencia elementos en *cis* necesarios para la replicación, debido a su proximidad con el extremo 3' del genoma, y para estudiar la actividad de los replicones generados mediante el análisis de la expresión del gen 7 por RT-PCR. Adicionalmente, se introdujo un sitio múltiple de clonaje conteniendo las dianas para las enzimas de corte único *AvrII*, *MluI*, *SwaI* y *FseI* a continuación del gen de la replicasa para permitir el clonaje y expresión de genes heterólogos (Fig. 25, detallado en Fig. 9).

2.5. ESTUDIO DE LA ACTIVIDAD DE LOS REPLICONES EN DIFERENTES LÍNEAS CELULARES

La actividad de los replicones se analizó en células humanas 293T y HeLa, células de hámster BHK-pAPN y células porcinas ST, que presentan eficiencias de transfección aproximadas del 42%, 18%, 20% y 5%, respectivamente. Los replicones REP 1, REP 2 y REP 3 se transfectaron con *Lipofectamine 2000* utilizando las condiciones optimizadas. En paralelo se transfectó la construcción pBAC-TGEV-NR-EGFP para medir en cada experimento la eficiencia de transfección y servir como control negativo de la actividad de la replicasa. Como control positivo se transfectó el clon infeccioso TGEV-RS del que se partió para la construcción de los replicones (Fig. 26A). Después de 40 h de la transfección de los distintos cDNAs, se extrajo el RNA total intracelular y se analizó la actividad de los replicones mediante la detección del mensajero subgenómico del gen 7 por RT-PCR utilizando los oligonucleótidos SP-VS y 7(213)RS (Tabla VIII) (Fig. 26A y B). En el caso de las células 293T y BHK-pAPN se detectaron niveles altos de mensajero del gen 7 a partir de la transfección de los replicones REP 1 y REP 2, indicando que estos eran funcionales en ambos tipos celulares, al menos a nivel de transcripción. Sin embargo, no se detectó mensajero en el caso del REP 3, sugiriendo que la proteína N, cuyo gen está deletado en esta construcción, podría ser necesaria para la actividad de los replicones. En el caso de las células ST y HeLa no se detectó actividad de ninguno de los replicones, posiblemente por la baja eficiencia de transfección de las células ST y la elevada citotoxicidad que produjo la transfección en el caso de las células HeLa. Para aumentar la sensibilidad del análisis, el mensajero del gen 7 se cuantificó mediante Q-RT-PCR utilizando los oligonucleótidos LDVS y 7RS (Tabla VIII). Se observó una diferencia de al menos tres unidades logarítmicas en la actividad de los replicones en las células 293T y BHK-pAPN en comparación con los niveles de actividad en las células ST y HeLa, lo que explica que no fueran detectables por RT-PCR convencional (Fig. 26C). También se detectó un nivel basal de actividad en el replicón REP 3, que era significativo respecto del control no replicativo, aunque muy inferior, entre dos y tres unidades logarítmicas, al de los replicones REP 2 y REP 1, respectivamente, indicando que la ausencia del gen N afectaba drásticamente a la actividad funcional del replicón REP 3. La diferencia de actividad entre los replicones REP 1 y REP 2 fue de una unidad logarítmica, superior en el caso del replicón REP 1, lo que podría deberse a que, en este replicón, la TRS del gen N se encuentra en su entorno de secuencia natural (Fig. 9).

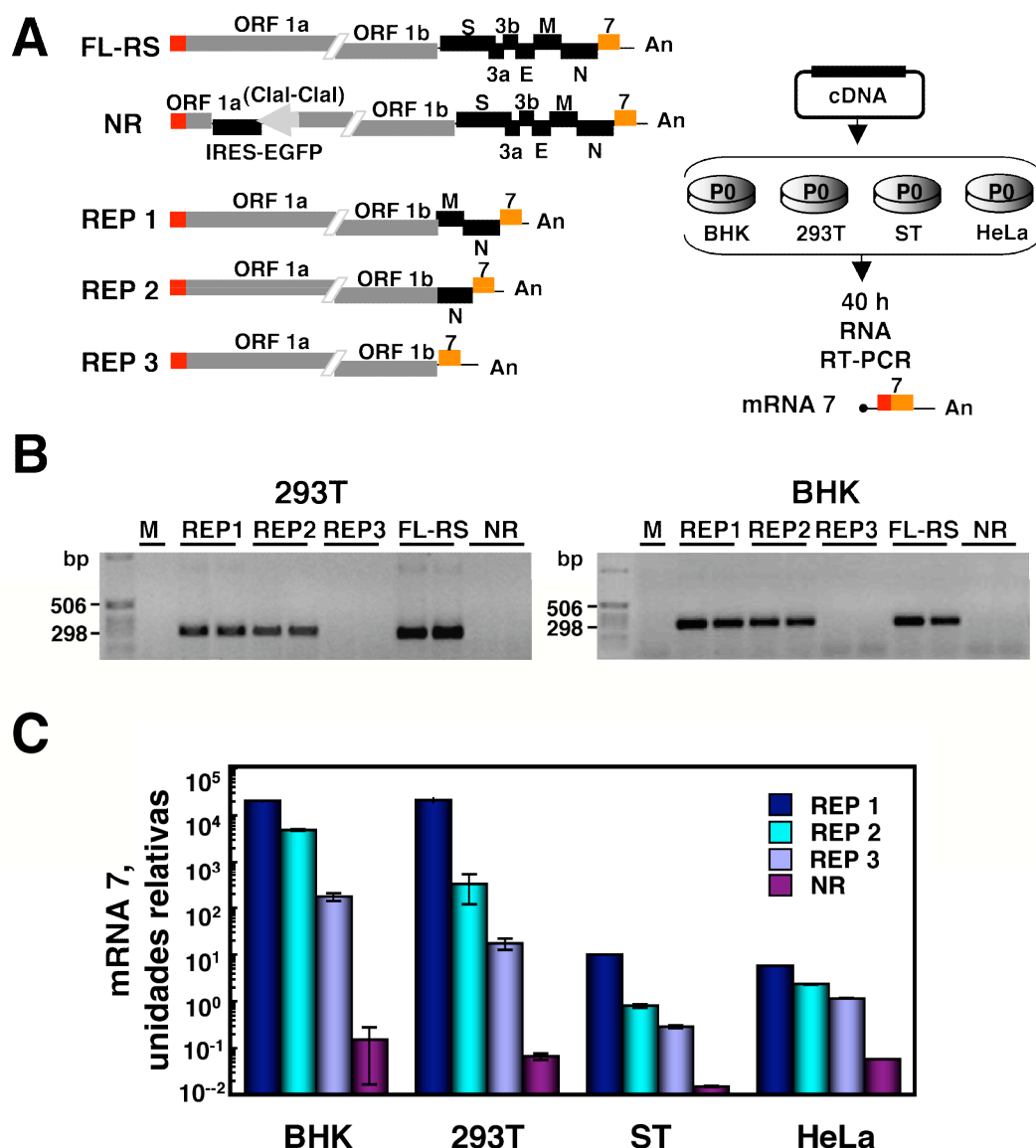


Figura 26. Actividad de los replicones del TGEV en células humanas 293T y HeLa, células porcinas ST y células de hámster BHK-pAPN. (A) Estructura genética de las construcciones transfectadas (izquierda) y diseño experimental (derecha). La actividad de los replicones REP 1 REP 2 y REP 3 del TGEV se analizó en células 293T, HeLa, ST y BHK-pAPN, llevando como control positivo el cDNA de TGEV que se utilizó de partida para construir los replicones (FL-RS) y un clon no replicativo del TGEV como control negativo (NR). A las 40 hdt se recogió el RNA para analizar la actividad de los replicones, mediante la detección por RT-PCR del mensajero del gen 7. En cada construcción se indican los genes virales. An, poliA. El fragmento *ClaI-ClaI* clonado en la orientación invertida se representa con una flecha en el control no replicativo (NR). (B) RT-PCR. Detección del mensajero del gen 7 en células 293T y BHK-pAPN transfectadas con los distintos replicones. En las células HeLa y ST no se detectó actividad mediante RT-PCR convencional. (C) Q-RT-PCR. La cuantificación del mensajero del gen 7 se evaluó mediante Q-RT-PCR en los distintos tipos celulares y se expresó en unidades relativas. La desviación estándar de dos cuantificaciones independientes se representa con barras de error.

2.6. LOS REPLICONES DERIVADOS DEL TGEV NO SON CITOPÁTICOS EN CÉLULAS HUMANAS 293T

Durante el análisis de la actividad de los distintos replicones en el pase P0, se observó un efecto citopático claro en el caso de las células HeLa y ST que no permitió el análisis de la actividad de los replicones a tiempos posteriores a las 40 hdt (datos no mostrados). Sin embargo, en el caso de las células humanas 293T, no se observó citotoxicidad, por lo que las células se pasaron sucesivamente mediante diluciones de una sexta parte del cultivo inicial. En cada pase se analizó el porcentaje de células transfectadas mediante citometría de flujo, utilizando para ello el control transfectado con la construcción pBAC-TGEV-NR-EGFP, y se determinó la actividad de los tres replicones mediante RT-PCR (Fig. 27A). El porcentaje de células transfectadas se analizó hasta el pase P3, observándose una reducción en un factor similar a la dilución equivalente del cultivo, sin observar efecto citopático aparente. En el caso de las células transfectadas con los replicones tampoco se detectó un efecto citopático claro, aunque en el pase P0 se detectó un cierto retraso en el crecimiento de las células transfectadas con las construcciones REP 1, REP 2 y REP 3 en relación con el control sin transfectar o las células transfectadas con el control no replicativo. Este retraso en el crecimiento

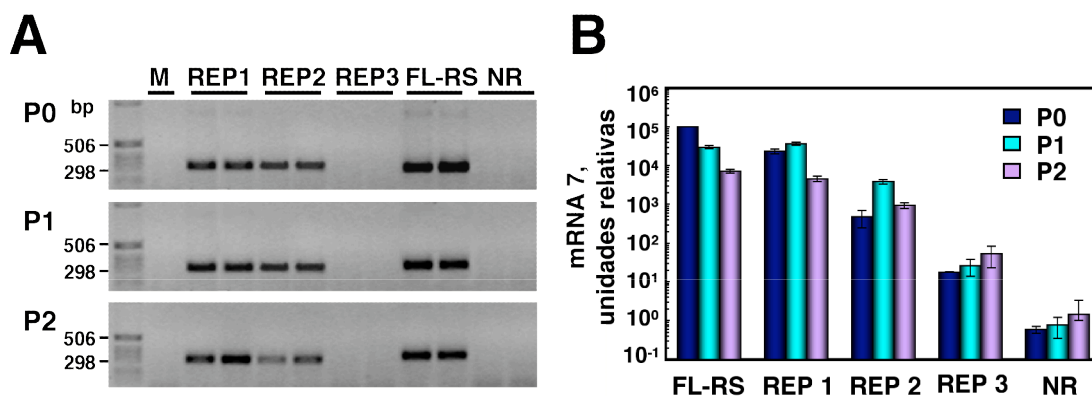


Figura 27. Actividad de replicones del TGEV a lo largo de los pases en cultivo. (A) RT-PCR. Detección del mensajero del gen 7 en células 293T transfectadas con los distintos replicones a lo largo de los pases. (C) Q-RT-PCR. Determinación de la actividad de los replicones mediante Q-RT-PCR a lo largo de los pases. La cuantificación del mensajero del gen 7 está expresada en unidades relativas y se indica la desviación estándar de dos cuantificaciones independientes con barras de error.

fue más acusado en las células transfectadas con el cDNA infectivo, posiblemente debido a la formación y liberación de virus infectivo en este caso (datos no mostrados). La actividad de los replicones a lo largo de los pases se determinó mediante RT-PCR (Fig. 27A) y se observó que la expresión del mensajero del gen 7 se mantenía con el

tiempo. Posteriormente, los niveles del mensajero del gen 7 se cuantificaron mediante Q-RT-PCR (Fig. 27B), lo que confirmó que la actividad de los replicones se mantenía a lo largo de los pases, y que disminuía en un factor correspondiente con la dilución del cultivo celular, indicando que los replicones no eran citopáticos en las células humanas 293T.

2.7. EXPRESIÓN DE GENES HETERÓLOGOS A PARTIR DE UN REPLICÓN DEL TGEV

Con el objetivo de evaluar la capacidad de los replicones derivados del TGEV para la expresión de genes heterólogos, se seleccionó el replicón REP 2, que contiene el gen de la proteína N y se insertó la secuencia que codifica la proteína EGFP precedida por la

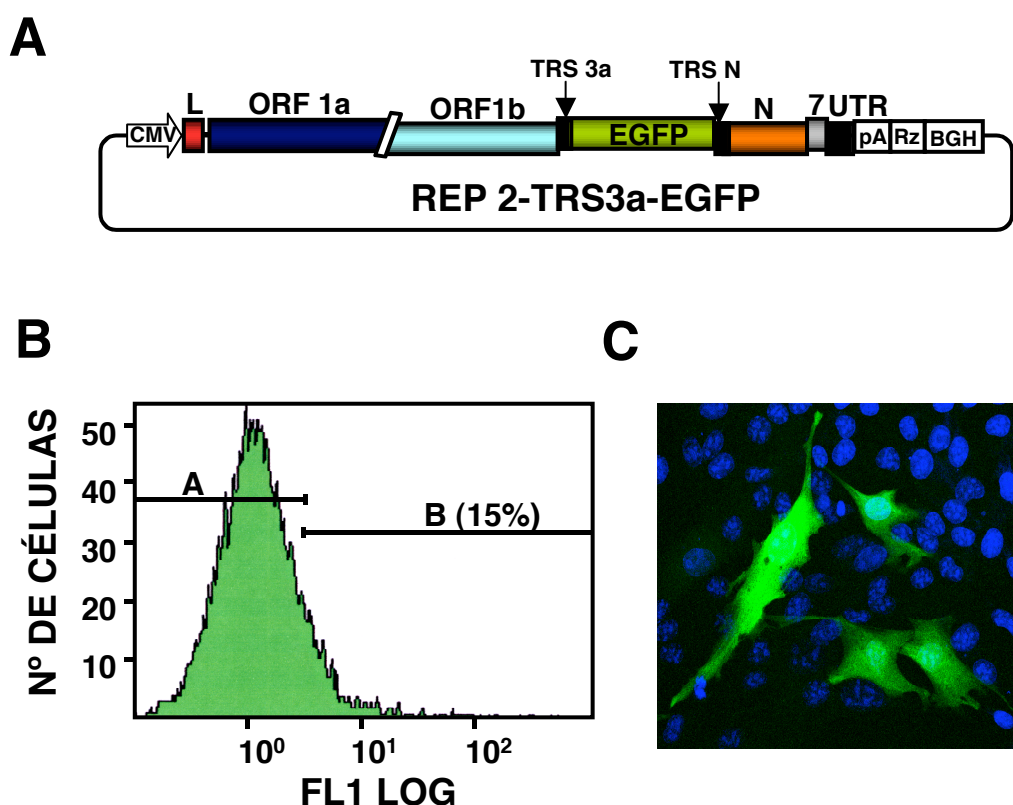


Figura 28. Expresión de genes heterólogos desde el replicón REP 2 del TGEV. (A) Estructura genética del replicón REP 2 del TGEV que expresa la proteína EGFP. El gen de la EGFP se introdujo bajo el control de la TRS del gen 3a. Se indican mediante flechas las posiciones de las TRSs. Se muestran los genes virales y secuencias flanqueantes. CMV, promotor de citomegalovirus; pA, poliA; Rz, secuencia de la ribozima del HDV; BGH, secuencia de terminación y poliadenilación del gen de la hormona BGH. (B) Evaluación de la expresión de la proteína EGFP desde el replicón REP 2 mediante citometría de flujo. Se indican con A y B las poblaciones de células negativas y positivas para la expresión de EGFP, respectivamente. La intensidad de fluorescencia, en unidades logarítmicas relativas, se representa en el eje de abscisas. (C) Análisis de la expresión de EGFP en células BHK-pAPN transfectadas con el replicón mediante microscopía de fluorescencia directa. Los núcleos se visualizaron mediante tinción con el reactivo TOPRO 3.

TRS del gen 3a (TRS3a), originando el replicón REP 2-TRS3a-EGFP (Fig. 28A). Esta construcción se transfectó en células BHK-pAPN con *Lipofectamine 2000* y a las 36 hdt la expresión de EGFP se analizó mediante citometría de flujo (Fig. 28B) y microscopía confocal (Fig. 28C). El 18% de las células se transfectaron eficientemente, de las cuales más del 80% expresaban niveles altos de EGFP, indicando que una mayoría de las células transfectadas expresaba el gen heterólogo.

3. ESTUDIO DEL PAPEL DE LA NUCLEOPROTEÍNA VIRAL EN LA SÍNTESIS DE RNA DE CORONAVIRUS

3.1. EFECTO DE LA PROTEÍNA N SOBRE LA ACTIVIDAD DE LOS REPLICONES

Los estudios de actividad de los replicones REP 1, REP 2 y REP 3 indicaban que la proteína N, cuyo gen está ausente en el replicón REP 3, podía ser la responsable de la actividad reducida de este replicón. Para analizar el papel de la proteína N sobre la actividad de los replicones, se utilizó un sistema de complementación en *trans* de la proteína, en el cual los replicones del TGEV se transfectaron en células BHK-pAPN previamente transfectadas con un replicón del virus Sindbis (pSIN-Rep21) que expresaba la proteína N del TGEV. Este sistema había sido anteriormente utilizado en el laboratorio para complementar la expresión de la proteína E del TGEV en un virus recombinante en el que se había eliminado su expresión (Ortego y col., 2002).

Para generar células BHK-pAPN que expresaran la proteína N, éstas se transfectaron con un replicón del virus Sindbis en el que se había clonado el gen N del TGEV bajo el control de un promotor de transcripción interno (pSIN-Rep21-N). Esta construcción expresa el gen *pac* de resistencia a la puromicina desde otro mensajero subgenómico, lo que permite seleccionar las células transfectadas con el replicón (Fig. 29A). En paralelo, se utilizaron como controles células BHK-pAPN transfectadas con el replicón del virus Sindbis que expresaba la proteína E del TGEV (Ortego y col., 2002), células transfectadas con el replicón pSIN-Rep21 vacío y células BHK-pAPN sin transfectar. Después de la selección de las células transfectadas con los replicones de Sindbis en presencia de puromicina, se analizó la expresión de las proteínas N y E mediante inmunofluorescencia, utilizando un anticuerpo monoclonal específico para la proteína N (3BD10) y un antisuero policlonal que reconoce la proteína E (S2), observándose altos niveles de expresión (Fig. 29B). A continuación se realizó el ensayo de complementación en *trans* con los replicones del TGEV. Para ello las distintas células transfectadas con los replicones de Sindbis se transfectaron con los replicones REP 1, REP 2 y REP 3 del TGEV. A las 36 hdt se recogió el RNA total y se analizó la actividad de los replicones mediante la detección del mensajero del gen 7 por RT-PCR utilizando los oligonucleótidos SP-VS y 7(213)RS (Tabla VIII). En todos los casos se detectaron altos niveles del mensajero del gen 7 para los replicones REP 1 y REP 2, indicando que la actividad del replicón de Sindbis no interfería con la de los replicones

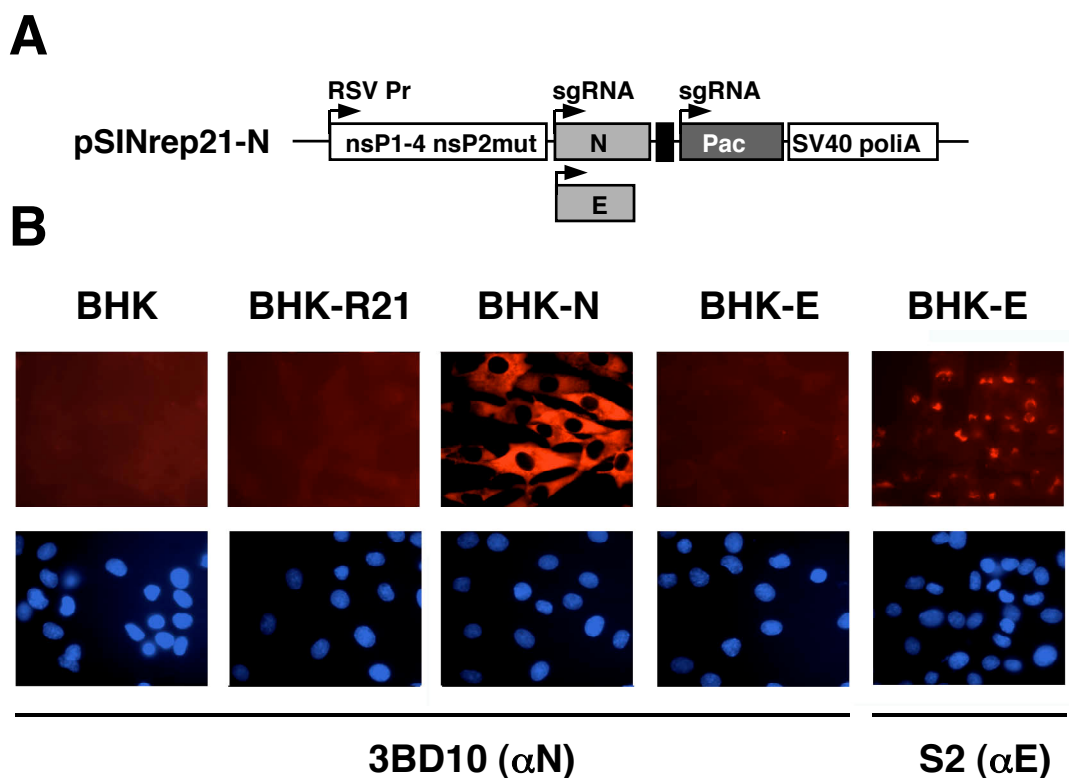


Figura 29. Expresión de las proteínas N y E del TGEV por el replicón del virus Sindbis en células BHK-pAPN. (A) Esquema del replicón de Sindbis como vector de expresión. Se representa el cDNA con el replicón no citopático de Sindbis que se utilizó para la expresión de la proteína N del TGEV desde un promotor subgenómico (pSINrep21-N). En paralelo se expresó la proteína E del TGEV en el mismo sistema de Sindbis y se utilizó como control. Los genes de la replicasa de Sindbis contenidos en el replicón codifican las cuatro proteínas no estructurales del extremo N-terminal (nsP1-4) conteniendo mutaciones adaptativas en la nsP2 (nsp2-mut). RSVPr, promotor del virus del sarcoma de Rous; N y E, genes N y E del TGEV; sgRNA, RNA subgenómico; Pac, gen de la puromicina N-acetiltransferasa; SV40 poliA, secuencia de terminación de la transcripción y poliadenilación del virus SV40. (B) Immunofluorescencia. Las células BHK-pAPN se transfectaron con los replicones de Sindbis conteniendo el gen N o E del TGEV y después de la selección con puromicina se analizó la expresión de las proteínas N y E mediante inmunofluorescencia utilizando los anticuerpos específicos 3BD10 (α N) y S2 (α E). En los paneles inferiores se muestra la tinción de los núcleos con el reactivo TOPRO 3.

del TGEV. En el caso del replicón REP 3, sólo se detectaron niveles altos del mRNA del gen 7 en aquellas células que expresaban la proteína N en *trans* (Fig. 30A), indicando que su expresión es necesaria para la actividad eficiente de este replicón. Para estudiar el grado de complementación de la proteína N aportada en *trans* sobre la actividad de los tres replicones, los niveles de mensajero del gen 7 se cuantificaron mediante Q-RT-PCR (Fig. 30B). En las células que no expresaban la proteína N se observó una actividad basal del replicón REP 3, que aumentó más de dos unidades logarítmicas cuando la proteína N estaba presente en *cis* (comparación entre los

replicones REP1 y REP 2 frente al REP 3 en células BHK-APN) y hasta tres unidades logarítmicas cuando se analizó en células que expresaban la proteína N en *trans* (comparación del replicón REP 3 en células BHK y BHK+N) (Fig. 30B). Por otra parte, la actividad de los replicones REP 1 y REP 2 también aumentaba en presencia de la proteína N expresada en *trans* desde el replicón de Sindbis, sugiriendo un efecto de dosis de la proteína N, o que su presencia fuera necesaria desde un principio.

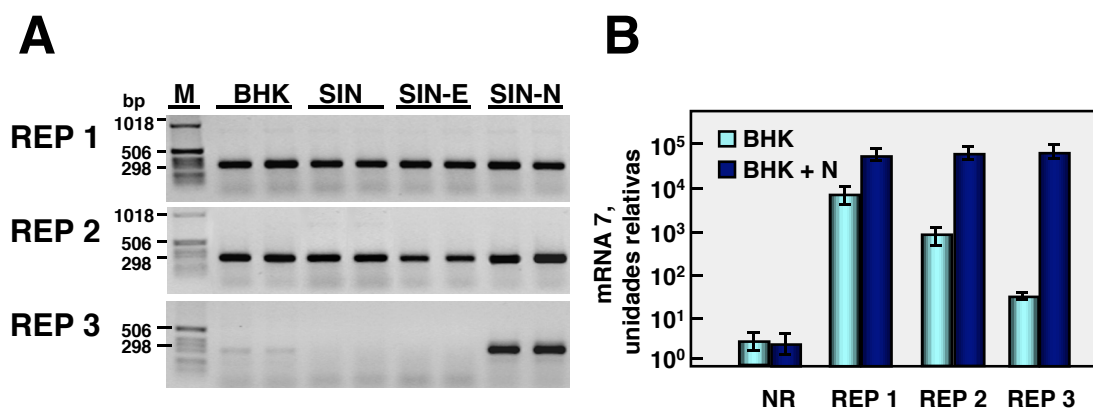


Figura 30. Efecto de la proteína N sobre la actividad de los replicones derivados del TGEV. (A) Análisis mediante RT-PCR del mensajero del gen 7. Las células BHK-pAPN (BHK) o BHK-pAPN transfectadas con el replicón de Sindbis vacío (SIN) o bien expresando la proteína E (SIN-E) o la proteína N del TGEV (SIN-N) se transfectaron con los replicones del TGEV (REP 1, 2 y 3) y a las 36 hdt se extrajo el RNA y se analizó mediante RT-PCR la expresión del gen 7. M, marcadores de masa molecular. Se muestra en cada caso el resultado de duplicados independientes. (B) Cuantificación mediante Q-RT-PCR del mensajero del gen 7. La cantidad de mensajero del gen 7, expresada en unidades relativas, se determinó mediante Q-RT-PCR utilizando los oligonucleótidos LDVS y 7RS (Tabla VIII). Se cuantificó el RNA aislado a las 36 hdt procedente de células BHK-pAPN (BHK) o células BHK-pAPN transformadas con el replicón de Sindbis, expresando la proteína N del TGEV (BHK + N), transfectadas con un control no replicativo del TGEV (NR) o bien con los replicones REP 1, REP 2 o REP 3. Se muestran los valores medios de tres experimentos, con las desviaciones estándar representadas como barras de error.

3.2. ESTUDIO DEL EFECTO DE LA PROTEÍNA N O SU mRNA SOBRE LA ACTIVIDAD DE LOS REPLICONES

Los resultados anteriores indicaban que la presencia de la proteína N era esencial para la actividad eficiente de los replicones del TGEV, aportada tanto en *cis* como en *trans*. Sin embargo, no se podía descartar que fuera el mRNA, y no la presencia de la proteína N, el responsable del aumento de la actividad de los replicones.

Para analizar esta posibilidad sin modificar la estructura genética de los replicones, se decidió realizar el experimento de complementación en *trans* con un replicón de Sindbis que expresara el gen de la N del TGEV con mutaciones puntuales que

impidieran la traducción de la proteína, y analizar su efecto sobre la actividad del replicón REP 3 del TGEV. Para ello se generaron tres mutantes del replicón de Sindbis en los cuales se mutaron una, dos o las tres metioninas presentes en la secuencia codificante del gen N del TGEV con el fin de suprimir la expresión de la proteína N completa o algún producto de iniciación interna de la traducción que se pudiera originar (Fig. 31A). Se generaron células BHK-pAPN expresando cada una de estas construcciones y se comprobó mediante inmunofluorescencia la ausencia de expresión de la proteína N (datos no mostrados). A continuación, el replicón REP 3 del TGEV se transfectó en paralelo en las células que expresaban la N en *trans* (N) o bien en las células que expresaban el RNA del gen N pero no la proteína (N1, N2, N3) (Fig. 31B). La actividad del replicón REP 3 se evaluó cuantificando el mensajero del gen 7

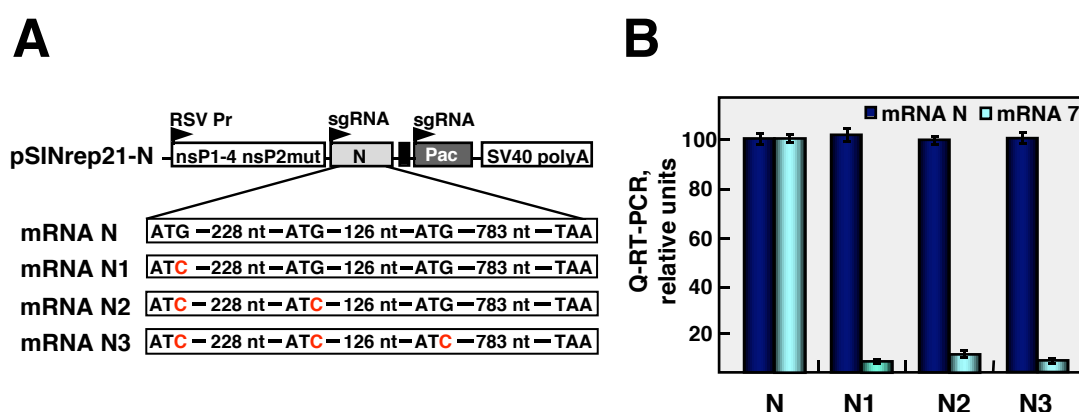


Figura 31. Efecto de la proteína N y su mRNA sobre la actividad de los replicones derivados del TGEV. (A) Esquema del replicón de Sindbis expresando el mRNA del gen N del TGEV o bien los distintos mRNAs mutantes. Las mutaciones introducidas en los tres ATGs de la secuencia codificante del gen N se indican en rojo. Las abreviaturas en el esquema del replicón de Sindbis se corresponden con las de la Fig. 29. (B) Cuantificación mediante Q-RT-PCR de los mRNAs de los genes N y 7. El Replicón REP 3 del TGEV se transfectó en células BHK-pAPN transfectadas con el replicón de Sindbis expresando el gen N sin mutar (N), o los mRNAs con mutaciones en el primer ATG del gen N (N1), los dos primeros ATGs (N2), o los tres ATGs (N3). Se cuantificaron mediante Q-RT-PCR los niveles de los distintos mRNAs del gen N expresados por el replicón de Sindbis, utilizando los oligonucleótidos Q1-N-VS y Q1-N-RS (Tabla VIII), y del mRNA del gen 7 expresado por el replicón REP 3 del TGEV, con los oligonucleótidos LDVS y 7RS (Tabla VIII). Se muestran los valores medios de dos experimentos, con las desviaciones estándar como barras de error.

mediante Q-RT-PCR. Adicionalmente, se cuantificaron los niveles del mensajero de la proteína N, expresados desde los replicones del virus Sindbis, que fueron similares en todos los casos (Fig. 31B). Se observó una clara reducción de la actividad del replicón REP 3 en las células transformadas con los mutantes del mensajero no traducible del

Resultados

gen N, en comparación con la actividad eficiente del replicón en las células que expresaban el mensajero y la proteína, lo que confirmaba que el aumento de actividad del replicón era debido a la actividad de la proteína N y no a la expresión de su mRNA.

4. IDENTIFICACIÓN DE PROTEÍNAS VIRALES Y CELULARES QUE INTERACCIONAN CON LOS EXTREMOS 5' Y 3' DEL GENOMA

En la replicación de coronavirus, además de las proteínas del gen de la replicasa, presumiblemente se incluyen proteínas celulares (Enjuanes y col., 2006; Ziebuhr, 2005). Asimismo, la proteína estructural N también desempeña un papel esencial en la replicación eficiente del coronavirus TGEV, como se demuestra en esta memoria. Las secuencias necesarias para la replicación del genoma de coronavirus se localizan en los extremos del genoma (Brian y Baric, 2005). El reconocimiento específico de estas secuencias por proteínas virales y celulares es probablemente esencial para la replicación, por lo que se planteó como objetivo la identificación de proteínas, virales y celulares que interaccionan con los extremos 5' y 3' del genoma. Con este fin, se diseñó una cromatografía de afinidad a RNA (Fig. 32), utilizando como ligando los extremos del genoma del TGEV marcados con biotina y extractos celulares citoplásmicos como fuente de proteínas. En un primer paso, los RNAs marcados con biotina se

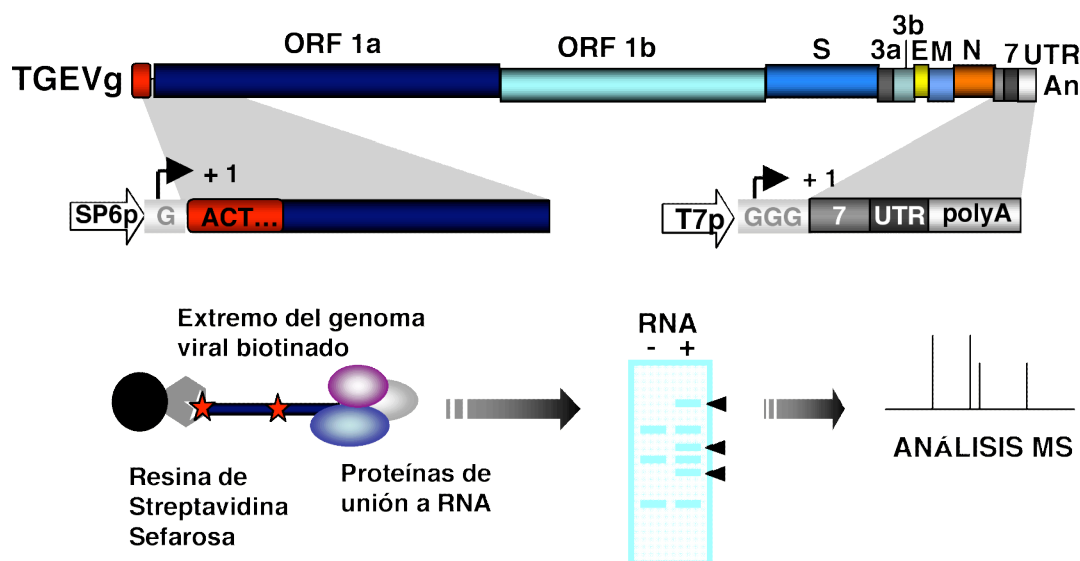


Figura 32. Estrategia para la identificación de proteínas que interaccionan con los extremos del genoma del TGEV. Esquema del genoma del TGEV (barra superior) indicando los genes virales, y de los moldes de transcripción *in vitro* de los extremos 5' y 3' (barras inferiores). En estas construcciones se indica el promotor de transcripción (T7p o SP6p, promotores de los fagos T7 o SP6, respectivamente), los nucleótidos no virales introducidos al inicio de la transcripción (cajas grises) y la posición del inicio de la transcripción (+1). En la parte inferior, se muestra el esquema de la cromatografía de afinidad de RNA para el aislamiento e identificación de proteínas que interaccionan con los extremos 5' y 3' del genoma. MS, espectrometría de masas.

inmovilizaron en una resina de sefarosa recubierta de estreptavidina mediante la interacción estreptavidina-biotina. A continuación el RNA inmovilizado se incubó con extractos citoplásmicos y los complejos de RNA-proteínas, se aislaron por centrifugación. Finalmente, las proteínas aisladas por afinidad al RNA biotinado se separaron mediante electroforesis monodimensional en paralelo con una muestra control sin RNA. Las bandas de interés se cortaron del gel, se digirieron con tripsina y se identificaron mediante espectrometría de masas.

4.1. GENERACIÓN DE RNAs BIOTINADOS DE LOS EXTREMOS DEL GENOMA VIRAL

Las secuencias en *cis* necesarias para la replicación de coronavirus residen en los extremos del genoma viral, y su extensión ha sido estudiada utilizando RNAs defectivos (Tabla II). En el caso del coronavirus TGEV, el minigenoma M33, de 3.3 kb, es el minigenoma más pequeño que se rescata con alta eficiencia (Izeta y col., 1999). Este minigenoma está formado por cuatro discontinuidades de secuencia respecto del genoma viral y contiene los 2144 primeros nucleótidos continuos del extremo 5' del genoma y los 495 últimos nucleótidos del extremo 3'. Un estudio reciente realizado en el laboratorio con mutantes de delección contruidos a partir de este minigenoma han mostrado que la secuencia necesaria para su amplificación eficiente en el extremo 5' está comprendida entre 1 y 1,6 kb (C. Capiscol, datos no publicados), aunque en otros coronavirus del grupo 2 como el MHV-A59 esta secuencia ha sido acotada hasta 466 nt (Luytjes y col., 1996).

Teniendo en cuenta estos estudios funcionales, para analizar la unión de proteínas a los extremos del virus se generaron distintos RNAs del extremo 5' del genoma de 0.5, 1, y 1.6 kb de longitud y uno del extremo 3' de 0.5 kb. Para ello se construyeron moldes de transcripción *in vitro* a partir del cDNA del minigenoma M33L. Se evitó modificar, en lo posible, la secuencia de los extremos del virus, cuya terminación precisa se ha descrito que es crítica para la replicación eficiente de otros virus (Herold y Andino, 2000; Silvestri y col., 2006), utilizando distintas estrategias (Fig. 33A). El extremo 5' del genoma se clonó bajo el promotor del fago SP6, en el que se introdujo la modificación de una única guanina ocupando la posición +1 del inicio de la transcripción. Para la generación de un RNA con el extremo 3' preciso del genoma, se construyó un molde de transcripción bajo el promotor del fago T7 que se linealizó

después de la secuencia de poliA con la enzima *HindIII*. Esta enzima dejaba un extremo 5' protuberante en la cadena que actúa de molde para la transcripción *in vitro*, por lo que se romizó posteriormente mediante incubación con la nucleasa de cadena sencilla *Mung Bean*, para generar transcritos con el extremo preciso de poliA.

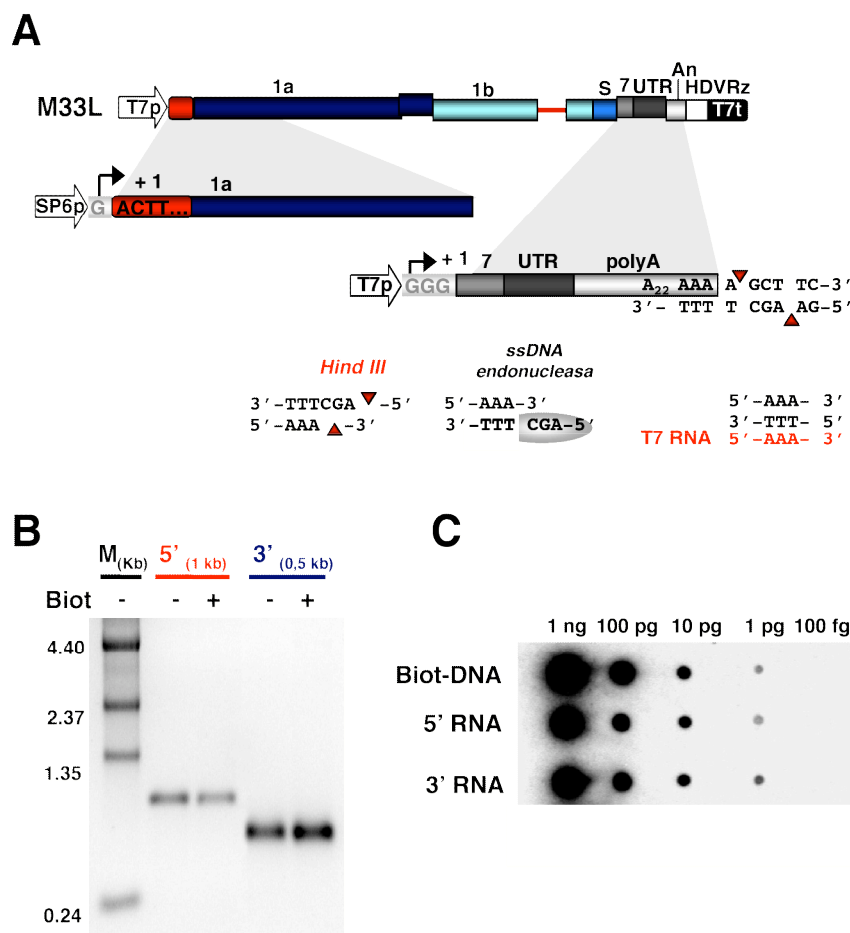


Figura 33. Generación de RNAs biotinados de los extremos 5' y 3' del genoma del TGEV. (A) Construcción de moldes de transcripción *in vitro* a partir del minigenoma M33L. SP6p y T7p, promotores de los fagos SP6 y T7, respectivamente. (+1), posición del inicio de la transcripción. Para generar un RNA del extremo 3' terminado en un poliA el molde se linealizó con la enzima de restricción *HindIII* y se romizó con la endonucleasa de cadena sencilla *Mung Bean*. (B) Análisis del tamaño e integridad de los transcritos. Los RNAs se visualizaron mediante tinción con bromuro de etidio en geles desnaturalizantes de 1% agarosa-2.2 M formaldehído. La movilidad de los RNAs marcados con biotina (Biot +) fue idéntica a los no marcados (Biot -). (C) Análisis del marcaje con biotina. El marcaje de los RNAs se analizó en un ensayo de *dot-blot* que se reveló mediante quimioluminiscencia, en comparación con un DNA biotinado como control positivo (Biot-DNA).

Para obtener RNAs biotinados de los extremos del genoma del TGEV, los moldes de DNA generados se transcribieron *in vitro* en presencia del análogo de ribonucleótido biotin-14-CTP. Este marcaje puede afectar a la conformación del RNA y a su

interacción con proteínas, por lo que se utilizó una relación molar de biotin-14-CTP de 1:5.25 con respecto al nucleótido no marcado, lo que permite el marcaje espaciado del RNA. Estas condiciones han sido optimizadas para la captura de proteínas de interacción con RNAs de las regiones 5' y 3' no traducidas de pestivirus (Isken y col., 2003). Las transcripciones *in vitro* se realizaron a escala preparativa y los RNAs resultantes se analizaron en un gel desnaturalizante de agarosa para comprobar su tamaño e integridad (Fig. 33B). La movilidad de los RNAs biotinados en el gel fue idéntica a los no marcados. A continuación se comprobó el marcaje del RNA mediante la detección de los grupos de biotina en un ensayo de *dot-blot*, utilizando para su revelado un conjugado de streptavidina acoplada a fosfatasa alcalina y un sustrato quimioluminiscente (Fig. 33C). El marcaje de las sondas de RNA con biotina permitió la detección de hasta 1 pg de RNA.

4.2 CROMATOGRAFÍA DE AFINIDAD PARA LA CAPTURA DE PROTEÍNAS QUE INTERACCIONAN CON EL RNA VIRAL

Para aislar e identificar proteínas que interaccionen con los extremos del genoma del TGEV mediante proteómica, inicialmente se utilizaron extractos comerciales de células HeLa como fuente de proteínas. Estos extractos se seleccionaron con el fin de utilizar una fuente de proteínas humanas, cuyas secuencias están completas en las bases de datos a diferencia de las proteínas del cerdo, el hospedador natural del TGEV. Las células HeLa permiten la replicación del TGEV cuando se introdujo en forma de cDNA y la replicación del coronavirus humano HCoV-229E, estrechamente relacionado con el TGEV, por lo que deben contener probablemente los factores necesarios para la replicación viral. Para realizar la cromatografía de afinidad de RNA inicialmente se inmovilizó 1 µg de RNA biotinado como ligando a 20 µl de resina de estreptavidina-sefarosa en presencia de alta concentración de sal, para favorecer el estiramiento de las moléculas de RNA y facilitar la unión del RNA a la matriz. A continuación, la resina se lavó con el tampón de unión de proteínas, para evitar que la alta concentración de sal interfiriera posteriormente en las interacciones RNA-proteína. Una vez inmovilizado el RNA a la resina se incubó con los extractos de proteínas en un agitador rotatorio durante 14 h y, posteriormente, los complejos formados se recuperaron mediante centrifugación de la matriz. Las proteínas aisladas se lavaron sucesivas veces y finalmente se resuspendieron en tampón de muestra, se separaron mediante

electroforesis monodimensional en geles comerciales de poliacrilamida y se tiñeron con azul de Coomasie. Inicialmente, no se observaron diferencias entre las proteínas aisladas en presencia o ausencia de RNA biotinado (datos no mostrados), indicando que en las condiciones ensayadas había un elevado número de proteínas que se unían inespecíficamente a la resina, e impedía apreciar diferencias debidas a la presencia del RNA. Para eliminar las interacciones inespecíficas con la matriz, los extractos de proteína se preaclaron antes de la incubación con el RNA inmovilizado, mediante pasos sucesivos de incubación en presencia de la resina de estreptavidina-sefarosa. Después de cada paso de preaclorado, las proteínas aisladas en ausencia de RNA biotinado, se analizaron en un gel y se observó que el fondo inespecífico se reducía sucesivamente después de cada paso de preaclorado, hasta alcanzar un nivel mínimo

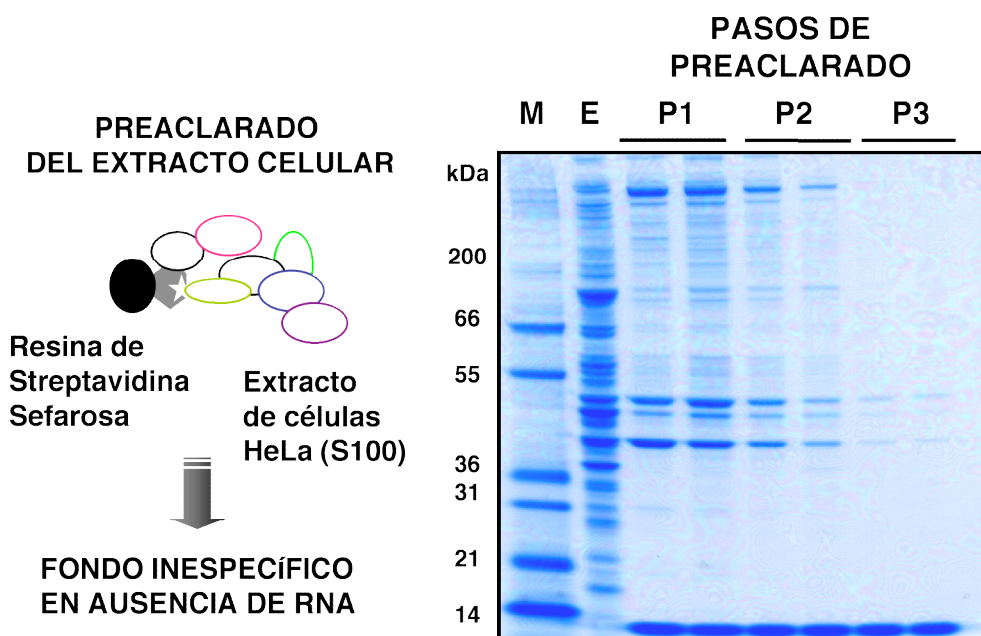


Figura 34. Reducción de fondo inespecífico de la cromatografía de afinidad a RNA. Para reducir el fondo inespecífico debido a la unión de proteínas a la matriz de estreptavidina-sefarosa se realizaron varios pasos de preaclorado de los extractos en presencia de la resina (esquema izquierdo). Las proteínas unidas a la resina en cada paso se resolvieron mediante SDS-PAGE y se visualizaron por tinción con azul de Coomasie. E, extracto de proteínas inicial; P1, P2 y P3, proteínas retenidas a lo largo de tres pasos de preaclorado. Los marcadores de masa molecular (M) se muestran a la izquierda en kDa.

después de tres pasos (Fig. 34). La adición de pasos adicionales de preaclorado no redujo significativamente este fondo.

Una vez reducido el fondo de unión inespecífica de las proteínas a la matriz, se realizaron cromatografías de afinidad de RNA utilizando como ligando los extremos 3'

(0,5 kb) y 5'(1 y 1,6 kb) del genoma viral. La estrategia de preaclara los extractos podría estar eliminando algunas proteínas de interés, pero permitió detectar proteínas que se unían específicamente en presencia de los RNAs biotinados, en comparación con la resina control sin RNA (datos no mostrados). La eficacia del ensayo fue menor en el caso de utilizar RNA de mayor tamaño, debido probablemente a un mayor impedimento estérico para su inmovilización a la resina o su interacción con las proteínas. En un primer análisis, las proteínas aisladas en presencia del RNA se recortaron del gel, se digirieron con tripsina y se analizaron mediante cromatografía líquida en fase reversa y espectrometría de masas (LC/MS/MS). Se consideraron identificaciones fiables las proteínas que presentaron una puntuación estadísticamente significativa asignada por el programa de búsqueda MASCOT, si además se obtenía un mínimo de dos péptidos secuenciados de la proteína identificada y la proteína correspondía por especie de origen y masa molecular, a las características de la banda recortada. Como resultado de este primer análisis, se identificó un conjunto de proteínas, en su mayoría de unión a RNA, tales como la proteína del RNA heterogéneo nuclear U (hnRNP U), la ribonucleoproteína La, o el factor de unión a la caja Y (YB-1) que se unieron a los dos extremos del genoma. Otras proteínas tales como la helicasa de RNA DHX15 o el factor nuclear de 45 kDa (NF45) se unían sólo al extremo 3'. Estas diferencias se consideraron provisionales debido a que la eficiencia de la cromatografía fue diferente según el tamaño de los RNAs. Otras proteínas como la proteína de unión a tramos de poliadenina (PABP) o la hnRNPQ, que se unían al extremo 3', fueron asignadas con un grado de significancia bajo. Sin embargo, ambas proteínas se habían identificado anteriormente por su interacción con los extremos del genoma del coronavirus MHV (Choi y col., 2004; Spagnolo y Hogue, 2000). Para la obtención de datos reproducibles se decidió optimizar la unión de proteínas al RNA de los extremos 5' y 3' del genoma. Asimismo, para igualar la eficacia y especificidad de la cromatografía de afinidad de RNA, se realizaron los ensayos sucesivos con RNAs biotinados del mismo tamaño (0.5 kb).

Con el fin de mejorar la eficiencia del ensayo, se aumentó la dosis de RNA biotinado. Para ello se realizó la cromatografía de afinidad de RNA utilizado como ligando de 2 a 10 µg de RNA y se analizaron las proteínas unidas en cada caso (Fig. 35). Se observó un aumento significativo de la sensibilidad del ensayo con la dosis de RNA, por lo que se decidió utilizar 10 µg de RNA biotinado como ligando en los experimentos siguientes.

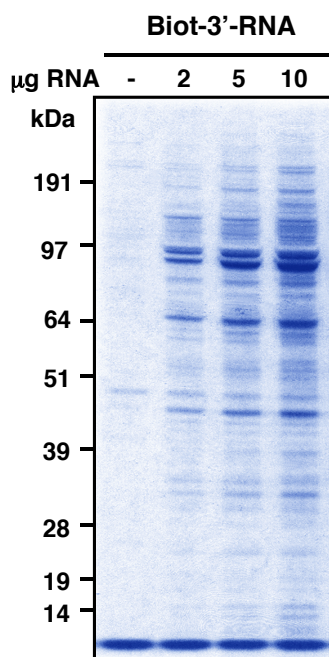


Figura 35. Análisis de la cantidad de RNA biotinado utilizada para el aislamiento de proteínas de unión a los extremos del genoma. Para aumentar el rendimiento de la cromatografía de afinidad de RNA, se probaron diferentes dosis de RNA biotinado del extremo 3' del genoma (Biot-3'-RNA). El RNA biotinado se inmovilizó en la resina de estreptavidina-sefarosa y se incubó posteriormente con extractos de células HeLa, anteriormente preacclarados en presencia de la resina. Después de tres lavados las proteínas unidas se eluyeron, se analizaron mediante SDS-PAGE y se visualizaron por tinción con Coomassie.

4.3. IDENTIFICACIÓN DE PROTEÍNAS QUE INTERACCIONAN CON LOS EXTREMOS DEL GENOMA

Una vez optimizado el ensayo para el aislamiento e identificación de proteínas de unión a RNA, se realizaron cromatografías de afinidad utilizando como ligando secuencias RNA de los extremos del virus, de 0.5 kb de longitud. Como fuente de proteínas se utilizaron extractos citoplásmicos de líneas celulares pertenecientes a tres especies distintas: células humanas HeLa, para facilitar las identificaciones en las bases de datos, células porcinas ST, procedentes del hospedador natural del virus y células BHK-pAPN, de hámster que junto con las células ST son susceptibles al TGEV. La utilización de las células BHK-pAPN podría suponer una ventaja en la identificación de proteínas mediante proteómica por homología con las de ratón, cuyas secuencias está mejor representadas en las bases de datos que las proteínas de cerdo.

Las cromatografías de afinidad de RNA se realizaron por duplicado, utilizando extractos citoplásmicos de los tres tipos celulares y 10 μ g de RNA biotinado de cada uno de los extremos del virus. Las proteínas aisladas por su afinidad con los extremos 5' y 3' del genoma mostraron un patrón de movilidad muy similar (Fig. 36). Interesantemente, en algunos casos se observaron proteínas de unión preferencial a uno de los extremos y algunas de estas interacciones se reprodujeron en los extractos celulares de las tres especies (Fig. 36). La mayor parte de las proteínas de los tres tipos

celulares aisladas por interacción con los extremos 5' y 3' que estaban ausentes en el control sin RNA se recortaron para su identificación. En los casos en que se observaban proteínas diferencialmente unidas a uno de los extremos o con distinta afinidad, se recortó la banda correspondiente a la proteína diferencial y la porción de gel de tamaño equivalente en el carril correspondiente a las proteínas de unión con el otro extremo del genoma. Las bandas de proteína se digirieron con tripsina en gel y se identificaron por espectrometría de masas. Una vez obtenida su huella peptídica, se realizaron búsquedas automatizadas en la base de datos del NCBI utilizando el programa MASCOT. Se identificaron un total de 40 proteínas entre los distintos tipos celulares, de las que se seleccionaron 30 que se identificaron un mínimo de dos veces con una puntuación estadísticamente significativa en los diferentes experimentos (Tablas IX y X). En todos los casos se comprobó si el tamaño de la proteína identificada coincidía con el peso molecular de la banda de procedencia. El porcentaje de identificaciones varió en función de la especie de procedencia de las proteínas, desde un 99% en el caso de las células humanas, a un 50% en las células porcinas o un 45% en las células de hámster. Las proteínas de cerdo y de hámster, se identificaron en su mayoría por homología con las proteínas de otras especies, preferentemente humana, murina y bovina.

Del total de las proteínas identificadas, un 80% correspondieron a proteínas previamente descritas por su unión a RNA. El resto de proteínas identificadas podrían estar copurificando con las de unión a RNA a través de interacciones proteína-proteína o bien podrían ser ribonucleoproteínas no caracterizadas. En las células ST y BHK-pAPN infectadas se identificó claramente la nucleoproteína viral que, además de ser estructural, es esencial en la replicación de coronavirus, como se ha descrito en los resultados de esta tesis.

Una de las proteínas aisladas por afinidad al extremo 3' del genoma en extractos de células ST infectadas presentaba el tamaño esperado para la RdRp viral (100 kDa). El análisis de esta proteína mediante espectrometría de masas MALDI-TOF permitió identificar varios péptidos tripticos que se correspondían con la polimerasa del virus, aunque la asignación se obtuvo con un grado de significancia bajo (resultados no mostrados) y necesita su confirmación inequívoca mediante la secuenciación de los péptidos obtenidos. En el laboratorio, la RdRp se ha identificado significativamente por su unión a la secuencia complementaria al líder, utilizando una aproximación similar (estudio en colaboración con la Dra Isabel Sola y Jose L. Moreno, datos no mostrados).

Un subconjunto de las proteínas, tales como la RNA helicasa A, la cadena pesada de la miosina o la hnRNP A1 se identificaron por su unión a RNA de forma reproducible con los extractos de proteína de las tres especies, coincidiendo con proteínas altamente conservadas, mientras que otras fueron identificadas sólo en alguna de las especies. Las diferencias observadas entre los tres tipos celulares podrían deberse a distintos factores, tales como el distinto grado de conservación de la secuencia según la especie, las diferencias en el proteoma de los distintos tipos celulares, o al grado de susceptibilidad al virus, en el caso de las células infectadas.

La mayor parte de las proteínas se unían a los dos extremos del genoma del TGEV (Tabla IX). No se observaron proteínas que se unieran preferentemente al extremo 5'

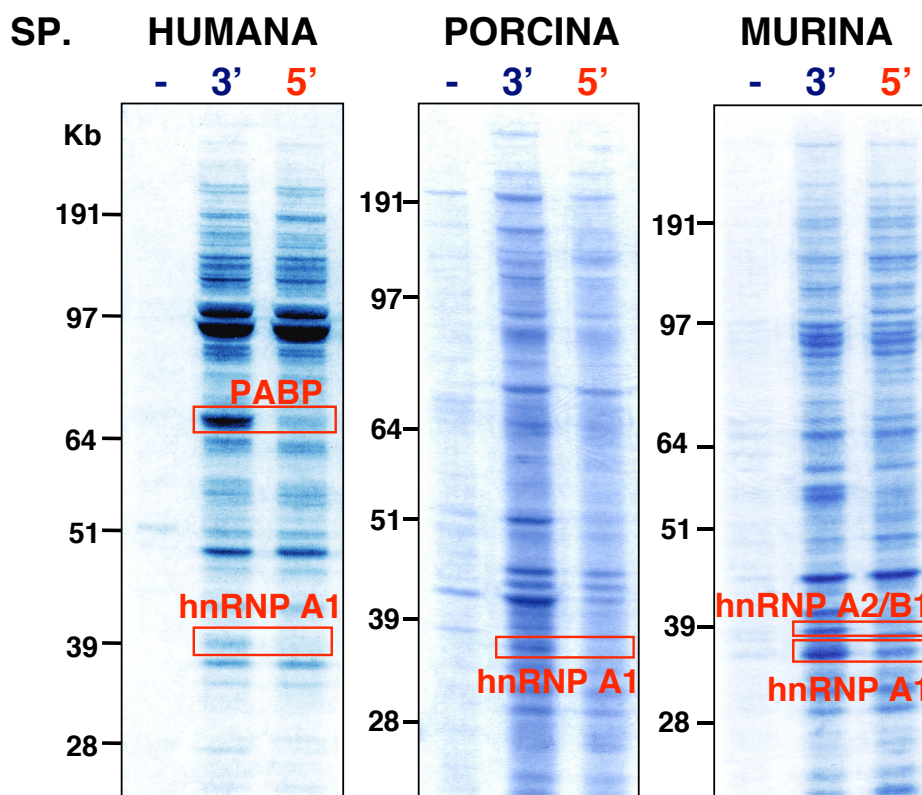


Figura 36. Aislamiento e identificación de proteínas que interaccionan con los extremos 5' y 3' del TGEV en extractos celulares de tres especies distintas. Se realizaron cromatografías de afinidad de RNA utilizando RNAs biotinados representando los extremos 5' y 3' del genoma viral y extractos de proteínas de tres especies susceptibles de soportar la replicación viral: extractos comerciales de células HeLa (SP. HUMANA), extractos citoplásmicos de células ST (SP. PORCINA) o de células de hámster BHK-pAPN (SP. MURINA). Algunas proteínas que mostraban una unión preferencial por el extremo 3' se indican con recuadros. La proteína hnRNP A1 se identificó en los extractos de las tres especies.

**TABLA IX. PROTEÍNAS QUE INTERACCIONAN CON LOS EXTREMOS 5' Y 3'
DEL GENOMA DEL TGEV**

CÉLULAS	RNA	N	[Nº Acceso] Identificación Especie (Masa molecular) Actividad/Función	PM/SS	MET
HeLa ST(i) BHK(i)	5' / 3'	6	[Q08211] RNA helicasa A dependiente de ATP <i>Homo sapiens</i> (140 kDa) Helicasa de RNA/DNA. Coactivador transcripcional	3/49/31	MS/MS ^a
HeLa ST(i)	5' / 3'	3	[Q12905] Factor de unión al enhancer de la IL-2, NF45 <i>Homo sapiens</i> (43 kDa) Unión a ácidos nucleicos de cadena doble. Factor de transcripción de la RNA pol II	2/95/47	MS/MS ^a
HeLa	5' / 3'	7	[P67809] Factor de transcripción de unión a la caja Y <i>Homo sapiens</i> (35.9 kDa) Unión a la caja-Y (CAAT) de RNA/DNA. Regulador transcripcional	6/250/45	MS/MS ^a
HeLa	5' / 3'	4	[P05455] Ribonucleoproteína La <i>Homo sapiens</i> (46.8 kDa) Factor de transcripción de la RNAPol III. Transporte de mRNA. Factor de traducción	4/132/45	MS/MS ^a
HeLa	5' / 3'	2	[Q6P2E9] Autoantígeno RCD8 <i>Homo sapiens</i> (151 kDa) Activador de la eliminación de cap	5/108/45	MS/MS ^a
HeLa ST(i)	5' / 3'	3	[P09405] Nucleolina <i>Mus musculus</i> (76,5 kDa) Une RNA. Transcripción del pre-rRNA y ensamblaje del ribosoma	14/153/78	HP ^a
ST(i)	5' / 3'	2	[P62270] Proteína ribosomal S18 <i>Mus musculus</i> (17 kDa) Proteína ribosomal	2/76/46	MS/MS ^a
HeLa ST(i)	5' / 3'	4	[P07814] Aminoacil-tRNA-sintetasa bifuncional <i>Homo sapiens</i> (163 kDa) Une RNA. Glutamyl-Prolyl-tRNA sintetasa	25/269/78	HP ^a
HeLa	5' / 3'	2	[O75534] Proteína con un dominio de choque térmico por frío E1, CSDE1 <i>Homo sapiens</i> (88 kDa) Une RNA. Regulador de la estabilidad del mRNA acoplada a la traducción	18/135/78	HP ^a
HeLa ST(i)	5' / 3'	6	[Q00839] hnRNP U <i>Homo sapiens</i> (90.4 kDa) Une ácidos nucleicos. Transcripción. Procesamiento del pre-mRNA	2/75/31	MS/MS ^a
HeLa ST(i)	5' / 3'	2	[O60506] hnRNP Q <i>Homo sapiens</i> (69.6 kDa) Procesamiento del mRNA	10/80/69	HP ^b
HeLa	5' / 3'	2	[P52272] hnRNP M, isoforma 2 <i>Homo sapiens</i> (77 kDa) Procesamiento del pre-mRNA	12/70/64	HP ^c
HeLa	5' / 3'	4	[O43143] Helicasa de RNA dependiente de ATP DHX15 <i>Homo sapiens</i> (90.9 kDa) Helicasa de RNA. Procesamiento del pre-mRNA. Desensamblaje de spliceosomas	4/132/45	MS/MS ^a
HeLa	5' / 3'	3	[O75533] Subunidad 1 del Factor de procesamiento de RNA 3B, SF3B 1 <i>Homo sapiens</i> (145.8 kDa) Componente del spliceosoma implicado en el ensamblaje de los complejos A y E	3/80/45	MS/MS ^a
BHK(i)	5' / 3'	2	[gi/82801284] Similar a Proteína heterogénea nuclear A3 <i>Mus musculus</i> (39 kDa) Procesamiento del pre-mRNA. Transporte de RNA en el citoplasma	12/108/78	HP ^b
HeLa BHK(i)	5' / 3'	3	[Q6P2Q9] Factor de procesamiento del RNA Prp8 <i>Homo sapiens</i> (220 kDa) Procesamiento del pre-mRNA	38/202/77	HP ^a
HeLa	5' / 3'	3	[O43172] Ribonucleoproteína pequeña nuclear U4/U6, Prp4 <i>Homo sapiens</i> (58.4 kDa) Procesamiento del pre-mRNA	12/101/78	HP ^a
HeLa	5' / 3'	2	[O43395] Factor de procesamiento del pre-mRNA 3, Prp3 <i>Homo sapiens</i> (77 kDa) Procesamiento del pre-mRNA	11/84/69	HP ^b
HeLa	5' / 3'	2	[O75643] Ribonucleoproteína pequeña nuclear U5, helicasa de 200 kDa <i>Homo sapiens</i> (200 kDa) Actividad helicasa de RNA dependiente de ATP	43/296/77	HP ^a
HeLa	5' / 3'	2	[Q99628] Proteína 1 de unión a Siah [Fragmento] <i>Homo sapiens</i> (57 kDa) Contiene un dominio de unión a poli(U) RNA	11/78/64	HP ^c
HeLa ST(i) BHK(i)	5' / 3'	2	[P35579] Cadena pesada de la miosina tipo A no muscular <i>Homo sapiens</i> (226 kDa) Une ATP. Actividad motora	9/211/45	MS/MS ^a
HeLa	5' / 3'	2	[Q4KMR3] Miosina IE <i>Homo sapiens</i> (127 kDa) Une ATP. Actividad motora	19/117/77	HP ^a
ST(i)	5' / 3'	2	[gi/1353212] Vimentina <i>Rattus norvegicus</i> (53 kDa) Citoesqueleto, filamento intermedio de clase III	17/159/69	HP ^b
ST(i)	5' / 3'	2	[gi/52631987] Annexin A2 <i>Sus scrofa</i> (38 kDa) Proteína de membrana regulada por calcio. Une RNA	14/140/69	HP ^b
HeLa	5' / 3'	3	[P11142] Proteína inducida por choque térmico de 70 kDa, hsp70 <i>Homo sapiens</i> (70,8 kDa) Actividad ATPasa. Chaperona molecular	11/100/78	HP ^a
HeLa ST(i)	5' / 3'	3	[P08238] Proteína de choque térmico de 90 kDa <i>Homo sapiens</i> (83.1 kDa) Chaperona molecular. Actividad ATPasa	7/166/31	MS/MS ^a
BHK(i) ST(i)	5' / 3'	6	[P05991] Proteína de la nucleocápsida <i>Coronavirus TGEV</i> (43 kDa) Estructural. Síntesis RNA viral	7/233/47	MS/MS ^a

TABLA X. PROTEÍNAS QUE INTERACCIONAN PREFERENTEMENTE CON EL EXTREMO 3' DEL GENOMA DEL TGEV

CÉLULAS	RNA	N	[Nº Acceso] Identificación Especie (Masa molecular) Actividad/Función	PM/S/SS	MÉT
BHK(i) HeLa ST(i)	3'	3	[gil57106316] Proteína similar a la hnRNP A1 <i>Canis familiaris</i> (34 kDa) Procesamiento alternativo y transporte núcleo-citoplasma del mRNA	4/66/62	HP ^c
BHK(i) ST(i)	3'	4	[gi/73976124] Similar a Proteína heterogénea nuclear A2/B1 <i>Canis familiaris</i> (32 kDa) Procesamiento del pre-mRNA	7/166/32	HP ^b
HeLa	3'	3	[P11940] Proteína de unión secuencias de poliA, PABP 1 <i>Homo sapiens</i> (70.6 kDa) Unión al poli(A) del mRNA. Traducción. Regulación de la estabilidad del mRNA	12/69/51	HP ^a

Tablas IX y X. Identificación de proteínas que interaccionan con los extremos 5' y 3' del genoma del TGEV. En las tablas IX y X se indica el origen de los extractos (CÉLULAS), el extremo de genoma con el que fue aislada (RNA), el número de veces que se reprodujo la identificación (N), el número de acceso de la proteína, la denominación, la especie de origen, la masa molecular teórica y la función. La utilización de células BHK-pAPN o ST infectadas se indica con la anotación (i). El número de acceso corresponde a la base de datos del SwissProt (formato P67809) o en su defecto del NCBI (formato gi/57106316). Los parámetros de los datos de proteómica se indican como: PM, número máximo de péptidos identificados; S, valor de probabilidad asignado por el programa de búsqueda MASCOT; SS, valor a partir del cual la identificación es estadísticamente significativa; MET, método de espectrometría de masas utilizado, bien mediante huella peptídica (HP) o mediante secuenciación de péptidos tripticos (MS/MS). Los superíndices indican la base de datos frente a la que se realizó la búsqueda: a, base del NCBI sin restricciones taxonómicas; b, base del NCBI, mamíferos; c, base del NCBI, humano.

del genoma. Sin embargo, algunas proteínas mostraron una unión preferente al extremo 3', como son la proteína PABP y distintas isoformas de la familia de las hnRNPs, como la hnRNPA1 y la hnRNPA2/B1 y se identificaron reproduciblemente en las células de las tres especies (Tabla X). Las proteínas PABP y hnRNP A1 han sido identificadas recientemente por su interacción con el genoma del virus BCoV, perteneciente al grupo 3 de coronavirus (Spagnolo y Hogue, 2000) y al MHV (grupo 2) (Furuya y Lai, 1993; Li y col., 1997; Shi y col., 2000; Zhang y Lai, 1995), respectivamente.

Las proteínas identificadas por su interacción con los extremos del genoma (Tablas IX y X) se clasificaron según su función celular (Fig. 37A). Mayoritariamente, las proteínas identificadas están implicadas en el metabolismo del RNA, regulando su transcripción, procesamiento, traducción y estabilidad. La clasificación de estas proteínas celulares dentro de estas categorías supone una simplificación, debido a que estos procesos están física y funcionalmente conectados y a menudo comparten parte de su maquinaria (Iborra y col., 2004). Algunas proteínas que forman parte del citoesqueleto o son chaperonas moleculares podrían estar copurificando mediante interacciones proteína-proteína y no como consecuencia de una unión directa al RNA. Del total de proteínas identificadas, el 60% han sido previamente identificadas por su

Resultados

interacción con el genoma u otros componentes de diferentes virus, tanto DNA como RNA, incluyendo los coronavirus (Fig. 37B). Las restantes proteínas (40%) no se habían identificado en relación con el ciclo de replicación de ningún virus, constituyendo un conjunto de candidatos nuevos, de interés para el estudio de las interacciones virus-huésped.

Interesantemente, algunas de las proteínas identificadas han sido implicadas en mediar la interacción entre los extremos del genoma de virus RNA. Tal es el caso de la PABP, que media la interacción 5'-3' del genoma de poliovirus (Herold y Andino, 2001) o miembros de la familia NFAR, de proteínas de unión a RNA de cadena doble (*Nuclear Factor Associated with double stranded RNA*), que incluyen a la RNA helicasa A y el factor NF45, y que probablemente median la interacción entre los extremos del genoma del pestivirus de la diarrea viral bovina (BVDV) (Isken y col., 2003).

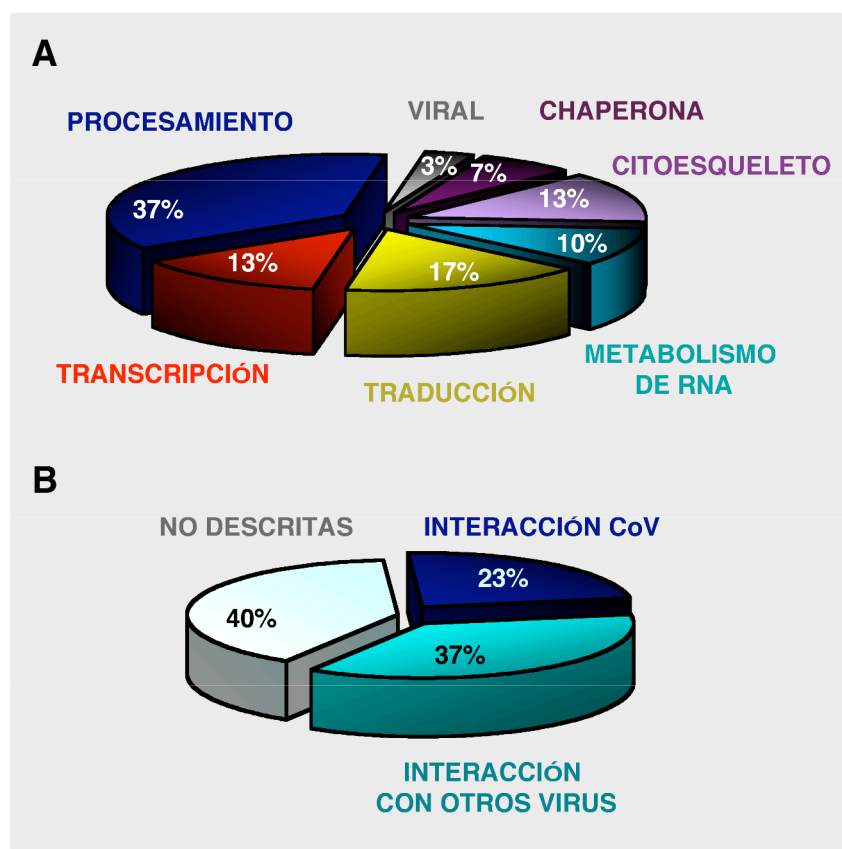


Figura 37. Clasificación funcional de las proteínas identificadas por su unión a los extremos del genoma del TGEV. (A) Clasificación atendiendo a su función celular, indicando en cada caso el porcentaje de proteínas identificadas con una función similar. (B) Clasificación atendiendo a su interacción con distintos virus. Las proteínas se clasificaron según su interacción descrita con componentes (genoma o proteínas) de otros coronavirus, o de otros virus RNA o DNA.

ESTUDIO DE LA INTERACCIÓN ENTRE LOS EXTREMOS DEL GENOMA VIRAL

La interacción entre los extremos del genoma de virus RNA se ha descrito en un número creciente de virus. Esta interacción tiene implicaciones funcionales en la traducción y replicación de estos virus, y puede estar mediada por proteínas o por hibridación de secuencias de RNA entre los dos extremos del genoma. En coronavirus existen datos funcionales que sugieren una interacción entre los extremos 5' y 3' del genoma durante la replicación, basados en el requerimiento de los dos extremos del RNA para la síntesis de la cadena positiva de minigenomas derivados de MHV (Lin y col., 1994). Sin embargo, los extremos del genoma de coronavirus no son complementarios, por lo que se ha postulado que esta interacción puede estar mediada por proteínas. Algunas de las proteínas identificadas por su interacción con los extremos del genoma del TGEV, como la PABP o las proteínas RNA helicasa A y NF45 del grupo NFAR, habían sido implicadas anteriormente en la interacción de los extremos 5' y 3' del genoma de virus RNA (Herold y Andino, 2001; Isken y col., 2003), por lo que podrían realizar una función similar en el virus TGEV. Para analizar esta hipótesis, se estudió la interacción *in vitro* entre los extremos 5' y 3' del genoma del TGEV en presencia o ausencia de proteínas aisladas por afinidad a los extremos.

4.2. DISEÑO DE UN ENSAYO DE COPRECIPITACIÓN DE RNA *IN VITRO*

Para estudiar la posible interacción entre los extremos del genoma del TGEV se puso a punto un ensayo de coprecipitación de RNA (Fig. 38A). Se sintetizaron fragmentos de RNA de ambos extremos del genoma del virus mediante transcripción *in vitro* y se marcaron de forma diferencial mediante la incorporación de grupos de biotina o digoxigenina en el RNA durante la transcripción. En un primer paso, el extremo marcado con biotina se inmovilizó a una resina de estreptavidina-sefarosa y se incubó a continuación con extractos celulares. Posteriormente los complejos RNA-proteína formados se aislaron por centrifugación y se añadió el otro extremo del genoma viral marcado con digoxigenina. Los complejos resultantes se volvieron a aislar por centrifugación y después de varios lavados se analizó la presencia del RNA marcado con digoxigenina coprecipitado en los complejos. Para ello, las proteínas unidas al RNA se eliminaron por tratamiento con proteinasa K, el RNA se resuspendió en tampón de muestra y se separó mediante electroforesis en un gel desnaturizante de 4%

poliacrilamida-8M urea, que se transfirió a una membrana de nailon. La coprecipitación del extremo de RNA marcado con digoxigenina se detectó específicamente mediante el revelado de la membrana con un anticuerpo anti-digoxigenina.

Para realizar este ensayo se utilizó el fragmento de RNA del extremo 3' de 0.5 kb que contiene todas las señales necesarias para la replicación de minigenomas y el fragmento de 1 kb del extremo 5', que contiene la mayor parte de las secuencias necesarias para la replicación. El extremo 3' se marcó en primer lugar con biotina y la presencia del extremo 5' marcado con digoxigenina se detectó en los complejos coprecipitados utilizando como fuente de proteínas extractos citoplásmicos de células ST infectadas (Fig. 38B). El ensayo de coprecipitación se reprodujo en el orden inverso, observándose la coprecipitación del extremo 3' marcado con digoxigenina. La presencia e integridad de los RNAs marcados con biotina en los complejos coprecipitados se

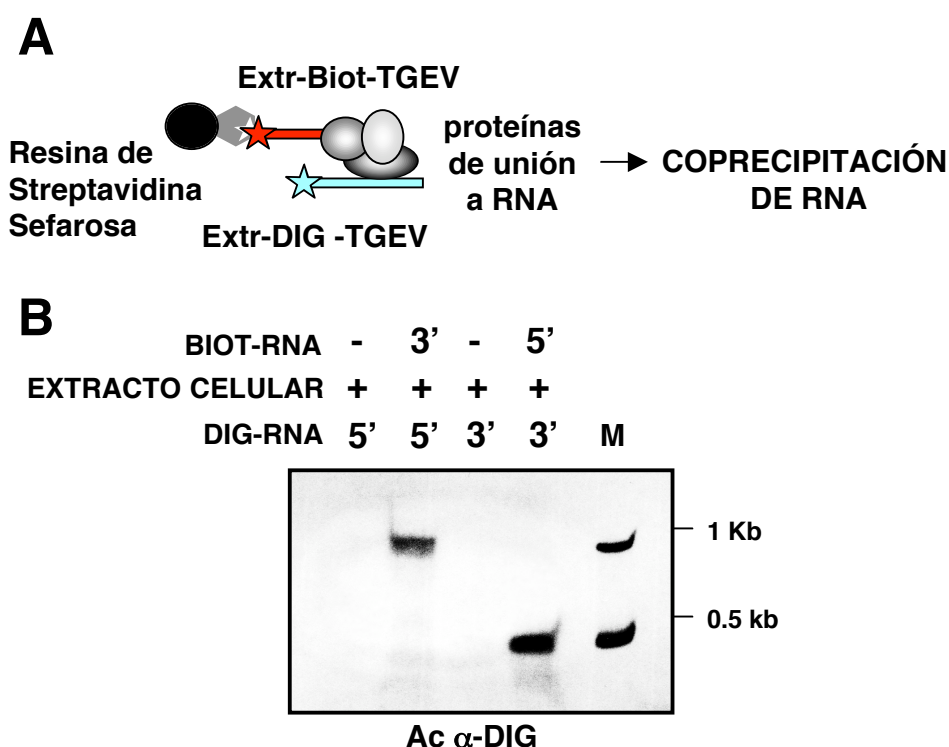


Figura 38. Coprecipitación del RNA de los extremos 5' y 3' del genoma en presencia de proteínas. (A) Esquema del diseño experimental. Los extremos 5' y 3' del genoma del TGEV se sintetizaron *in vitro* y se marcaron diferencialmente con biotina (Extr-Biot-TGEV) o digoxigenina (Extr-DIG-TGEV). El extremo de RNA biotinado se inmovilizó en una resina de estreptavidina-sefarosa y se incubó con extractos de células ST infectadas. Las proteínas aisladas se incubaron en un segundo paso con el RNA marcado con digoxigenina. Después de lavados sucesivos se analizó la coprecipitación del RNA marcado con digoxigenina. (B) Coprecipitación de RNA. El RNA se extrajo de los complejos con proteínas mediante tratamiento con proteinasa K y se analizó mediante electroforesis desnaturalizante, electrotransferecia a una membrana de nailon e inmunodetección con un anticuerpo específico acoplado a fosfatasa alcalina. M, transcritos de los extremos 5' y 3' marcados con digoxigenina, como marcador de masa molecular.

confirmó en un ensayo en paralelo, en el que la membrana se reveló con un conjugado de estreptavidina-fosfatasa alcalina y detección quimioluminiscente (resultados no mostrados). Los datos obtenidos indicaban que los extremos del genoma interactuaban *in vitro* y que esta interacción estaba mediada por proteínas.

4.3. ESTUDIO DE LOS REQUERIMIENTOS PARA LA INTERACCIÓN *IN VITRO* DE LOS EXTREMOS DEL GENOMA

Una vez observada la interacción *in vitro* entre los extremos del genoma, en presencia de proteínas procedentes de células ST infectadas, se estudió si la presencia de proteínas era necesaria para mediar la interacción y si, en su caso, eran necesarias proteínas virales o celulares. Para ello, el experimento de coprecipitación de RNA se realizó en presencia o ausencia de proteínas, procedentes de extractos de células ST infectadas o no infectadas. En ausencia de proteínas no se observó señal de coprecipitación entre los extremos 5' y 3' (Fig. 39) como era de esperar, puesto que los extremos del genoma del TGEV no son complementarios. Por el contrario, se observó coprecipitación de los extremos de RNA del genoma tanto en presencia de extractos de células infectadas como no infectadas. Estos resultados indicaron que las proteínas celulares eran suficientes para mediar la interacción de los extremos del TGEV.

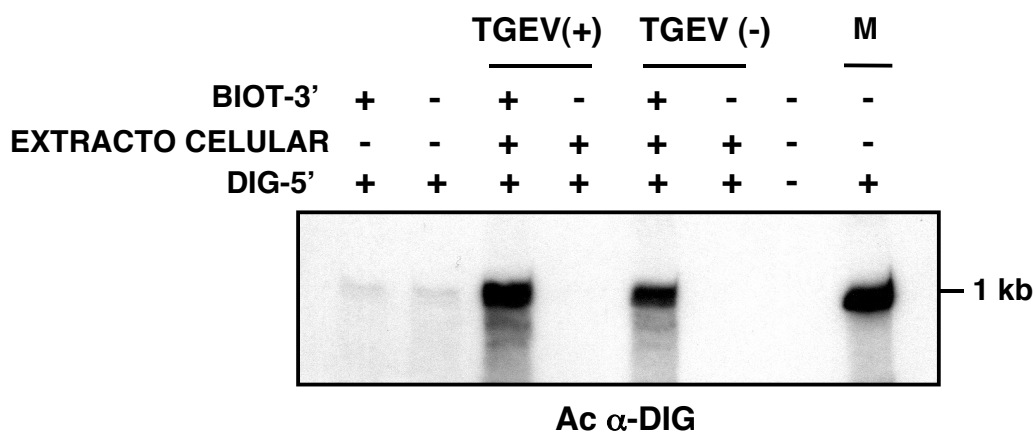


Figura 39. Requerimientos para la coprecipitación de los extremos 5' y 3' del genoma del TGEV. Se muestra la interacción entre el extremo 3' del genoma biotinado (BIOT-3') y el 5' marcado con digoxigenina (DIG-5') en ausencia o en presencia de proteínas (EXTRACTO CELULAR +/-) procedentes de extractos de células ST infectadas [TGEV (+)] o no infectadas [TGEV (-)]. La autorradiografía muestra la especie de RNA coprecipitado (el 5', marcado con digoxigenina). M, transcrito del extremo 5', como marcador de masa molecular.

4.4. ESTUDIO DE LA ESPECIFICIDAD DE LA INTERACCIÓN *IN VITRO* DE LOS EXTREMOS DEL GENOMA

Para analizar si la interacción de los extremos del genoma viral requería proteínas celulares o podría ocurrir en presencia de proteínas no específicas, se realizaron las coprecipitaciones de RNA en presencia de albúmina sérica bovina. En este caso no se observó coprecipitación de RNA (resultados no mostrados).

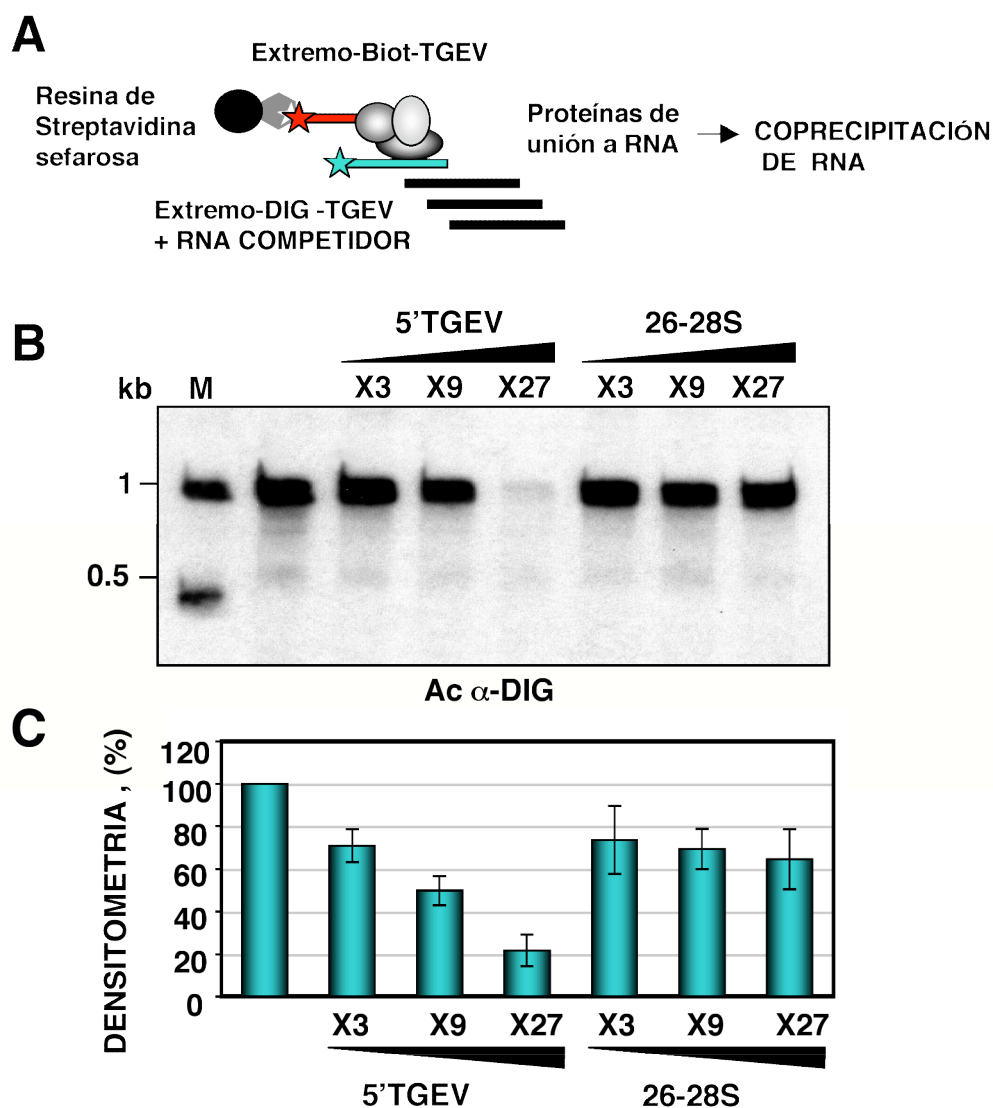


Figura 40. Competición de la interacción *in vitro* de los extremos 5' y 3' del genoma del TGEV. (A) Esquema del experimento de competición. El ensayo de coprecipitación de RNA se realizó en presencia o ausencia de un exceso de RNA competidor, que se añadió durante el paso de incubación con el RNA marcado con digoxigenina. (B) Coprecipitación de RNA. Autorradiografía del RNA coprecipitado marcado con digoxigenina (extremo 5') en ausencia o presencia de un competidor no marcado, homólogo (5' TGEV), o un RNA no relacionado (26-28S). Se indica el exceso de RNA competidor, en número de veces respecto al marcado. M, transcritos de los extremos 5' y 3' marcados con digoxigenina, como marcador de masa molecular. (C) Densitometría. La señal de coprecipitación de RNA se cuantificó por densitometría y los valores se expresaron en % con respecto a la señal de coprecipitación en ausencia de competidor, que se tomó como el 100%. Se muestran los valores medios de tres experimentos independientes, con las desviaciones estándar representadas como barras de error.

Para analizar la especificidad a nivel de RNA, el experimento de coprecipitación se realizó en presencia de un exceso de RNA no marcado como competidor (Fig. 40A). La competición con el RNA homólogo del extremo 5' redujo proporcionalmente la señal de coprecipitación del RNA (Fig. 40B) mientras que la competición con RNA ribosomal 26-28s no produjo una reducción significativa de la señal, indicando la existencia de un componente específico en la interacción.

En el momento de la escritura de esta memoria, se está analizando cuáles de las proteínas identificadas por su interacción con los extremos del genoma median la interacción 5'-3' de modo específico. Para ello, se está inhibiendo su expresión utilizando RNAs de interferencia para determinar su relevancia en la replicación viral. Las proteínas seleccionadas funcionalmente serán analizadas de forma individual para determinar cuáles de ellas median la interacción de los extremos 5' y 3' del genoma.

VII. DISCUSIÓN

En esta tesis se ha utilizado como modelo el virus TGEV para estudiar las bases moleculares de la replicación de coronavirus, un proceso regulado por secuencias en *cis* y factores en *trans*, entre los que se encuentran presumiblemente proteínas virales y celulares.

En primer lugar se ha estudiado una mutación puntual, que afecta al procesamiento *in vitro* nsp1/nsp2 de las poliproteínas de la replicasa, sobre la replicación del virus TGEV y la amplificación de minigenomas. Los requerimientos aparentes de este procesamiento fueron diferentes para la replicación del genoma y para la amplificación de los minigenomas derivados del TGEV, lo que sugiere posibles diferencias entre la replicación en *cis* y en *trans*.

Con el objetivo de poder evaluar la replicación viral en *cis* se ha generado una colección de replicones del TGEV, cuya actividad es dependiente de la proteína N viral, aportada tanto en *cis* como en *trans*, asignando por primera vez un papel en síntesis de RNA a una proteína estructural codificada fuera del gen de la replicasa. Asimismo, se ha demostrado que los replicones generados constituyen vectores virales bioseguros, con capacidad de expresión eficiente de genes heterólogos.

Con el fin de identificar los posibles componentes del complejo de replicación, se ha realizado por primera vez un análisis exhaustivo, mediante proteómica, de las proteínas virales y celulares que interaccionan con los extremos del genoma del TGEV, donde residen las señales de replicación. La interacción de proteínas celulares con los extremos del genoma podría mediar una comunicación 5'-3', según se ha observado en experimentos de coprecipitación *in vitro* de RNA.

1. UN CAMBIO DE AMINOÁCIDO EN LAS POLIPROTEÍNAS DE LA REPLICASA AFECTA DE FORMA DIFERENCIAL A LA REPLICACIÓN DEL VIRUS TGEV Y A LA AMPLIFICACIÓN DE MINIGENOMAS

En este estudio se ha caracterizado el efecto de la mutación puntual en la posición 637 del genoma del TGEV, localizada en el sitio de procesamiento N-terminal de las poliproteínas de la replicasa, sobre la replicación del virus y la amplificación de minigenomas. Para el rescate de virus a partir del clon infectivo y para la amplificación de minigenomas, se ha demostrado un requerimiento diferencial en la naturaleza del nucleótido en la posición 637. El procesamiento N-terminal de la replicasa por la

proteasa viral PLP1 se vio afectado por la mutación de la posición 637 que daba lugar a un cambio drástico de aminoácido, en un ensayo *in vitro* de procesamiento en *trans*. Las mutaciones que afectaban al procesamiento nsp1/nsp2 produjeron un defecto en el rescate de virus, sin embargo, fueron seleccionadas por los minigenomas para su amplificación eficiente.

La mutación puntual de la posición 637 del genoma se mantuvo originalmente en el clon infectivo del TGEV puesto que se encontraba presente en el RNA defectivo DI-C, que es amplificado eficientemente por el virus complementador y cuyo cDNA fue utilizado como base para ensamblar el clon infectivo (Almazán y col., 2000). A pesar de la relativa flexibilidad del clon infectivo del TGEV para su manipulación genética (Ortego y col., 2003), se observó una rápida reversión del marcador genético en la posición 637 del genoma en los virus recombinantes rescatados a partir del cDNA, lo que indicaba una fuerte presión de selección en esa posición.

La correlación entre el nucleótido 637 y el sitio de procesamiento nsp1/nsp2 de la replicasa del TGEV, se estableció por homología con el sitio de procesamiento equivalente de la replicasa del coronavirus humano HCoV-229E, que es procesado por la proteasa viral PLP1 (Herold y col., 1998). Los determinantes de reconocimiento del sustrato de las proteasas PLP1 de diferentes coronavirus, parecen ser dos residuos pequeños sin carga en las posiciones flanqueantes al sitio de corte (P1 y P1') y cierta preferencia de un residuo cargado positivamente en la posición P5 (Ziebuhr y col., 2000). De acuerdo con estos requerimientos, la enzima PLP1 del virus TGEV podría procesar el sitio nsp1/nsp2 en dos posiciones alternativas, entre los aminoácidos Thr107/Gly108 o Gly110/Ala111. Un cambio de nucleótido en la posición 637 afecta a la naturaleza del aminoácido 108, que ocuparía las posiciones P1' o P3 relativas a los sitios de procesamiento Thr107/Gly108 o Gly110/Ala111, respectivamente (Herold y col., 1998). En este trabajo, se ha demostrado que la introducción de un aminoácido no conservativo en la posición 108, afecta *in vitro* al procesamiento N-terminal de las poliproteínas de la replicasa mediado por la proteasa PLP1.

En un estudio reciente, se ha propuesto una nueva clasificación de las proteasas similares a papaína de coronavirus en dos grupos, en base a predicciones estructurales del sitio de unión al sustrato (Sulea y col., 2006). Un primer grupo de estas enzimas presentan un bolsillo de unión al sustrato similar a las proteasas específicas de ubiquitina, con restricción de tamaño en la posición P2 del sustrato a un residuo de Gly

o Ala (grupo-R, de *restricted*). El segundo grupo de proteasas presenta un sitio de unión al sustrato similar a la papaína, más abierto en la zona de interacción con el residuo P2 (grupo-O, de *open*). Según esta clasificación, la proteasa PLP1 del virus TGEV pertenece al grupo-R, y el sitio de procesamiento nsp1/nsp2, de acuerdo con la restricción de tamaño en la posición P2, coincidiría con la posición del sitio de corte equivalente del virus HCoV-229E (TGEV:Ala111/Ile112; HCoV-229E:G111/Asn112). En esta nueva posición teórica del sitio nsp1/nsp2, el aminoácido 108 ocuparía la posición P4 respecto del punto de corte, situada en la cavidad de reconocimiento del sustrato de la enzima, por lo que un cambio drástico de aminoácido produciría igualmente un defecto de procesamiento. Sin embargo, estas predicciones necesitan su comprobación experimental mediante la secuenciación N-terminal del sustrato procesado, que no ha podido ser llevada a cabo en nuestro estudio por la reducida actividad *in vitro* de la proteasa. La limitada actividad de la proteasa PLP-1 en el ensayo de procesamiento podría ser debida a la reducida longitud del sustrato (Teng y col., 1999), a la necesidad de dominios flanqueantes de la enzima necesarios para su actividad eficiente, o a un defecto de conformación de la enzima para mediar el procesamiento en *trans*, lo que está de acuerdo a lo observado anteriormente para la actividad en *trans* de diferentes proteasas de tipo PLP de coronavirus (Herold y col., 1998; Teng y col., 1999).

Las mutaciones en la posición 637 del clon infectivo del TGEV, que predeciblemente afectaban al procesamiento nsp1/nsp2, afectaban drásticamente al rescate de virus infectivo y al tamaño de las placas de lisis, lo que indica un papel esencial de las proteínas relacionadas con este sitio de procesamiento en la replicación eficiente del virus. En el coronavirus HCoV-229E, las proteínas N-terminales de las poliproteínas de la replicasa son un polipéptido de 9 kDa, o p9, que se corresponde con la nsp1 y una proteína de 87 kDa o p87, que se corresponde con la nsp2 (Herold y col., 1998). Hasta la fecha no se han obtenido datos funcionales para las proteínas nsp1 y nsp2 de coronavirus del grupo 1, y no se puede predecir su función en base a la comparación de su secuencia. Sin embargo, la liberación temprana de la proteína nsp1 observada durante la infección con el HCoV-229E (Herold y col., 1998), y el aparente requerimiento de este procesamiento para la producción viral del TGEV, sugieren que esta proteína juega un papel esencial en el establecimiento productivo de la infección.

La región N-terminal de la replicasa presenta gran divergencia en los diferentes coronavirus, cuyo significado biológico es desconocido. Así, las proteínas nsp1 (p28) y nsp2 (p65) de las poliproteínas de la replicasa de MHV, de 28 y 65 kDa, respectivamente, difieren en secuencia y tamaño con las proteínas nsp1 y nsp2 del TGEV. Las proteínas p28 y p65 de MHV colocalizan con los complejos de replicación del virus en membranas intracelulares (Brockway y col., 2004; Sims y col., 2000) y la expresión transitoria de la proteína p28 retrasa el crecimiento celular, induciendo la parada del ciclo celular en la fase G₀/G₁ (Chen y col., 2004). Esta parada del ciclo celular ocurre de forma temprana como consecuencia de la infección por MHV (Chen y Makino, 2004), aunque su relación causal con la proteína p28 ha de ser caracterizada posteriormente. Adicionalmente, el sitio de procesamiento nsp1/nsp2 se ha mutageneizado utilizando el clon infectivo del virus MHV-A59, y los datos obtenidos están de acuerdo con los observados en este trabajo, en cuanto a que las mutaciones que impedían este procesamiento afectaban drásticamente al crecimiento del virus y correlacionaban con placas de lisis de pequeño tamaño (Denison y col., 2004). Sin embargo, los mutantes de MHV en el sitio nsp1/nsp2 conservaban los cambios introducidos, a diferencia de lo observado en el virus TGEV, en el que las mutaciones no permisivas en la posición 637, revertían rápidamente, impidiendo la caracterización de los virus mutantes. Aunque los requerimientos de este sitio de procesamiento pudieran ser diferentes entre los virus TGEV y MHV por las divergencias específicas de cada grupo, parece claro que la mutación del sitio nsp1/nsp2 compromete drásticamente la viabilidad de ambos virus. Las proteínas nsp1 y nsp2 se ha observado que no son estrictamente esenciales, en su forma nativa, para la replicación de MHV, aunque distintos mutantes de delección produjeron un retraso en el crecimiento del virus o la reducción el título viral en cultivos celulares (Brockway y Denison, 2005; Graham y Denison, 2006; Graham y col., 2005).

En el caso del SARS-CoV, la proteína nsp1, de 20 kDa, es procesada por la única actividad proteasa del tipo PLP conservada en este virus y su expresión transitoria induce la degradación de los RNA mensajeros de la célula hospedadora (Kamitani y col., 2006), por lo que se ha propuesto como uno de los factores determinantes de la patogénesis del SARS-CoV. En un estudio independiente, se ha observado que en células humanas de pulmón que expresan la proteína nsp1 del SARS-CoV, ésta actúa como un potente inductor de expresión de citoquinas, lo que supondría un mecanismo

nuevo de este factor en patogénesis viral (Law y col., 2006). Sin embargo, para obtener conclusiones funcionales de la nsp1 del SARS-CoV será necesaria su caracterización en el contexto fisiológico de la infección.

A pesar de las diversas funciones propuestas para las proteínas nsp1 de los coronavirus del grupo 2 y su reducida homología de tamaño y secuencia con las del grupo 1, las proteínas de ambos grupos parecen conservar un dominio de dedo de zinc, que podría estar relacionado con su papel biológico (resultados no mostrados).

El efecto de las mutaciones en la posición 637 del genoma del TGEV se analizó en segundo lugar utilizando minigenomas. La obtención de minigenomas derivados del TGEV ha permitido acotar las secuencias en *cis* necesarias para su amplificación eficiente, evaluando su rescate a lo largo de los pases (Izeta y col., 1999). Sin embargo, el rescate de los minigenomas, necesita de su replicación y encapsidación eficientes, por lo que no es una medida directa de la replicación viral. Para analizar el efecto de las mutaciones en la posición 637 sobre la replicación de los minigenomas, se cuantificó inicialmente su acumulación en el pase P0 mediante Q-RT-PCR. Sin embargo, no se observaron diferencias significativas de replicación de los minigenomas en el pase P0, debido al bajo nivel de replicación, por lo que su amplificación tuvo que ser evaluada a lo largo de los pases. La falta de detección de la replicación en el pase P0 podría deberse a: (a) una baja eficiencia de replicación del minigenoma transfectado, debida a la modificación de su secuencia en el extremo 5', que comienza con tres "G" para favorecer su transcripción *in vitro*; (b) la degradación de parte de los minigenomas en la vía de entrada a través de lisosomas; (c) la localización del RNA transfectado fuera de los complejos de replicación virales; (d) el límite de sensibilidad de la Q-RT-PCR, capaz de distinguir diferencias de orden dos, que podría impedir apreciar la amplificación de los minigenomas respecto del total de RNA transfectado.

Al evaluar el rescate de los minigenomas a lo largo de los pases, se observó una fuerte dependencia del tipo celular, que mostró una eficiente replicación y producción viral del TGEV en las células porcinas ST, derivadas de su hospedador natural, en comparación con las células de hámster BHK-pAPN, pudiendo existir factores celulares implicados en estas diferencias.

El efecto de la mutación puntual en la posición 637 del genoma sobre la amplificación de RNAs defectivos se estudió finalmente evaluando el rescate de los minigenomas M33L y DI-C en células ST. Los requerimientos de secuencia en la

posición 637 para la replicación de los minigenomas fueron diferentes respecto de los observados para el rescate del virus TGEV. Sólo los minigenomas que presentaban mutaciones que predeciblemente impedían el procesamiento nsp1/nsp2 eran amplificados eficientemente por el virus complementador. Estas diferencias podrían tener implicaciones relevantes. Por un lado, las secuencias necesarias para la amplificación de los minigenomas, en *trans*, o del genoma viral, en *cis*, podrían ser diferentes, lo que cuestionaría la utilización de los RNAs defectivos para evaluar la replicación, favoreciendo la utilización de RNAs autoreplicativos o replicones en su lugar. Por otra parte, podrían existir diferentes estrategias para la replicación del genoma completo y de los minigenomas, que permitiría a los RNAs defectivos utilizar de forma ventajosa la maquinaria de replicación del virus complementador. Finalmente, los minigenomas utilizados para estos estudios, M33L y DI-C, codifican una proteína de fusión, cuyo procesamiento podría interferir específicamente con la amplificación de los

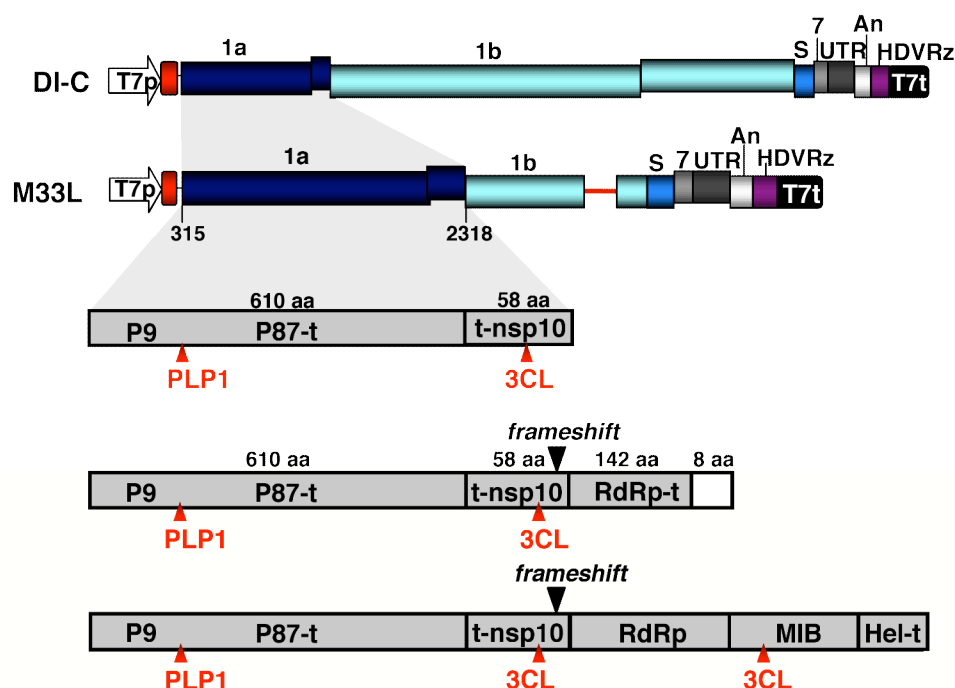


Figura 41. Proteínas de fusión codificadas por los minigenomas M33L y DI-C. Los minigenomas M33L y DI-C, se representan como barras en la parte superior, indicando los fragmentos de los genes virales y secuencias adyacentes. T7p y T7t, promotor y terminador del fago T7, respectivamente. UTR, región 3' no traducida; An, poliA; HDVRz, secuencia de la ribosoma del HDV. La proteína de fusión de la ORF1a común en los dos minigenomas se representa por una barra gris y se indica su procedencia con polígonos sombreados. En la parte inferior se indican las proteínas de fusión originadas por el salto de fase del ribosoma (*frameshift*) en los minigenomas M33L y DI-C (formas corta y larga, respectivamente). P87-t, polipéptido P87 truncado en su extremo C-terminal; t-nsp10, proteína nsp10 truncada en su extremo N-terminal; RdRp-t, RNA polimerasa dependiente de RNA truncada en su extremo C-terminal; MIB, dominio de unión a iones metálicos; Hel-t, helicasa truncada en su extremo C-terminal. Se indican los sitios de corte por las proteasas virales PLP1 y 3CL.

minigenomas. La traducción de esta proteína daría lugar a un producto común para ambos minigenomas, formado por secuencias fusionadas de la ORF1a. Alternativamente, podría ocurrir el salto de fase del ribosoma que produciría una versión extendida de la proteína de fusión para cada minigenoma (Fig. 41). Hasta la fecha no existe un consenso sobre los requerimientos de la traducción de los distintos RNAs defectivos de coronavirus caracterizados, aunque parece que la traducción, pero no el producto codificado, podría conferir una ventaja en la replicación de estos RNAs, probablemente mediante el acoplamiento de los mecanismos de traducción y replicación (Kim y col., 1993; Liao y Lai, 1995; van der Most y col., 1995). En el caso de los minigenomas M33L y DI-C, el procesamiento del sitio nsp1/nsp2 en la proteína de fusión liberaría la proteína terminal nsp1(p9) y un producto truncado, derivado de las proteínas nsp2-nsp10 (Fig. 41). Una interpretación posible de los resultados sería que esta proteína truncada actuara como dominante negativo, interfiriendo con la replicación de los minigenomas, impidiendo su acceso a la maquinaria de replicación del virus complementador, o formando complejos de replicación no funcionales.

Una explicación alternativa, teniendo en cuenta los datos observados con la nsp1 del SARS-CoV, cuya expresión parece promover la degradación de los mensajeros celulares, podría ser que la nsp1 procesada en los minigenomas, actuara degradando específicamente estos RNAs (Kamitani y col., 2006). Si este fuera el caso, el RNA del genoma de coronavirus debería estar específicamente protegido de este mecanismo de degradación, posiblemente localizándose en el microentorno de dobles membranas donde se localizan los complejos de replicación.

2. LA NUCLEOPROTEÍNA ES ESENCIAL PARA LA SÍNTESIS EFICIENTE DE RNA DE CORONAVIRUS

Hasta la reciente generación de clones infectivos de coronavirus, el estudio de su replicación a nivel molecular estaba restringido al análisis de RNAs defectivos (Izeta y col., 1999; Repass y Makino, 1998; Williams y col., 1999), de mutantes termosensibles deficientes en síntesis de RNA (Siddell y col., 2001; Stalcup y col., 1998), y de virus recombinantes generados mediante recombinación dirigida (Masters, 1999).

Los datos anteriores obtenidos en este trabajo planteaban algunas limitaciones en la utilización de RNAs defectivos para evaluar la replicación, debido a la dificultad de

detectar la síntesis de RNA en el pase P0. En este trabajo se describe la generación de una colección de replicones derivados del TGEV como una herramienta alternativa que permite evaluar la replicación en *cis* de forma independiente a las etapas de morfogénesis y ensamblaje de la partícula viral. Esta estrategia se ha utilizado anteriormente para estudiar la replicación y transcripción a nivel molecular de diferentes virus RNA de cadena positiva (Agapov y col., 1998; Lohmann y col., 1999; Molenkamp y col., 2000; Thiel y col., 2001b; Tijms y col., 2001). Para muchos de estos virus, sólo el gen de la replicasa es necesario para la replicación autónoma del RNA viral, incluyendo el caso del arterivirus equino EAV, perteneciente al orden *Nidovirales* (Molenkamp y col., 2000).

En una estrategia inicial se estudió la actividad de un replicón mínimo derivado del TGEV (TGEV-REP), que contenía únicamente el gen de la replicasa y los extremos 5' y 3' del genoma, aunque no se detectó actividad a nivel de transcripción o de amplificación en *trans* de minigenomas. Una de las limitaciones de este sistema era la baja eficiencia de transfección de los BAC en células ST (0,1%), que podía suponer un problema de sensibilidad a la hora de evaluar la actividad funcional del replicón. Alternativamente, podría ser necesario algún otro gen, además de la replicasa, para la amplificación eficiente de los replicones.

En primer lugar se optimizaron los niveles de transfección hasta obtener niveles del 20% en el caso de células BHK-pAPN y del 42% en el caso de las 293T. Esta optimización supuso un aumento de cinco unidades logarímicamente en el rescate de virus rTGEV a partir del clon infectivo, obteniéndose títulos de hasta 10^7 - 10^8 ufp/ml en el pase P0. El incremento en el rescate de virus supone una ventaja adicional del sistema de genética reversa basado en los BAC, con respecto a los basados en la ligación y transcripción *in vitro* del RNA viral (Yount y col., 2000; Yount y col., 2002), o los basados en el genoma de poxvirus (Casais y col., 2001; Thiel y col., 2001b).

Con el objetivo de analizar los requerimientos necesarios para la autoamplificación del RNA del TGEV, se eliminaron los genes 3a, 3b y 7, puesto que se había demostrado que no son esenciales para la replicación del virus en cultivos celulares, y de los genes E y S, que están implicados en las etapas de morfogénesis y no afectan a la replicación del RNA (Ortego y col., 2002; 2003; Sola y col., 2003). Se construyó una colección de tres replicones, mediante delección secuencial de los restantes genes estructurales del extremo 3' del genoma, y se demostró la necesidad de la proteína N, aportada en *cis* o

en *trans*, para la síntesis de RNA. En ausencia de la proteína N se pudo detectar un nivel basal de actividad (a partir del replicón 3) (Fig. 26), lo que está de acuerdo con los datos observados con un replicón basado en el coronavirus HCoV-229E, que carecía del gen N (Thiel y col., 2001a). Sin embargo, la actividad del replicón 3 aumentó 1000 veces en presencia de la proteína N, aportada en *trans* por el replicón de Sindbis, indicando que la actividad de la proteína N, aunque no es estrictamente esencial, es necesaria para la actividad eficiente de este replicón.

Los replicones 1 y 2, que contienen el gen N, presentaban diferencias en su actividad, siendo 10 veces superior en el caso del replicón 1. Esta diferencia cuantitativa probablemente está relacionada con diferentes niveles de expresión de la proteína N a partir de cada construcción. El replicón 1, que mantiene la secuencia del gen M en posición 5' relativa al gen N, contiene la secuencia TRS del gen N así como las secuencias flanqueantes de su entorno natural, mientras que en el replicón 2 se introdujo una TRS sintética del gen N de 22 nt, optimizada en un sistema de minigenomas defectivos (Alonso y col., 2002) que podría carecer de las secuencias adyacentes necesarias para la expresión óptima de este gen en el contexto del replicón (Fig. 9). En el ensayo de complementación en *trans* de la proteína N desde el replicón de Sindbis, la actividad de los replicones 1 y 2 aumentó 10 y 100 veces, respectivamente. Este efecto podría ser debido al aumento de la proteína aportada en *trans*, que incrementara la actividad de los replicones, o bien a que su presencia inicial en las células, anterior a la síntesis de la proteína por transcripción desde el replicón del TGEV, pudiera potenciar en mayor medida la actividad de los replicones (Fig. 30).

En estudios anteriores se había observado que la proteína N interaccionaba específicamente con secuencias implicadas en la regulación de la síntesis del RNA viral, como la secuencia líder, las TRSs y el extremo 3' del genoma (Baric y col., 1988; Cologna y col., 2000; Molenkamp y Spaan, 1997; Nelson y col., 2000; Stohlman y col., 1988), así como con la proteína hnRNP A1, recientemente implicada en la síntesis de RNA de coronavirus (Shi y col., 2000). Adicionalmente, la proteína N se concentra en los complejos de replicación a tiempos tempranos de la infección (Denison y col., 1999; Sims y col., 2000; van der Meer y col., 1999) y la cotransfección de transcritos de la proteína N mejora el rescate de virus infeccioso en diferentes sistemas de genética reversa (Casais y col., 2001; Yount y col., 2000; Yount y col., 2002). Aunque las observaciones anteriores sugieren que la proteína N puede tener un papel en la

replicación de coronavirus, el estudio presentado en este trabajo, establece por primera vez una correlación entre la expresión de la nucleoproteína y la síntesis de RNA viral. Este es el primer ejemplo descrito de una proteína estructural implicada en la síntesis de RNA de un virus de cadena positiva no segmentado. Entre todos los virus RNA de cadena positiva, sólo un grupo de virus de plantas pertenecientes a la familia *Bromoviridae*, que contienen un genoma segmentado, presentan un fenómeno similar (Bol, 1999; Jaspars, 1999; van der Vossen y col., 1994). Sin embargo, los resultados obtenidos no permiten determinar a qué nivel es necesaria la proteína N, pudiendo tener un papel en replicación o transcripción y ejercer una actividad directa o indirecta sobre la síntesis de RNA, a través de la interacción con otras proteínas o de protección del RNA una vez sintetizado.

Por otro lado, el requerimiento de la proteína N para la síntesis de RNA en el TGEV, podría diferir de lo observado en el caso del arterivirus EAV (Molenkamp y col., 2000). Esta diferencia, junto con el tamaño notablemente superior del genoma de los coronavirus y la presencia de actividades adicionales en el gen de la replicasa, indican que el mecanismo de replicación de los genomas de los nidovirales denominados “grandes” (Gorbalenya y col., 2006) presenta un grado de complejidad mayor, posiblemente orientado a mantener la estabilidad de estos genomas.

En este estudio la actividad de los replicones ha sido cuantificada a nivel de transcripción, debido a que es un sistema más específico para distinguir la síntesis debida a la RdRp viral respecto de la acumulación del RNA del replicón, a la que también contribuye la RNAPol II celular, cuando el replicón se transcribe desde un promotor de CMV. Recientemente, se ha generado un replicón del SARS-CoV en nuestro laboratorio siguiendo la misma estrategia, y se ha demostrado su actividad mediante Q-RT-PCR a nivel de replicación y transcripción (Almazán y col., 2006).

Hasta la fecha sólo se ha desarrollado otro replicón de coronavirus, basado en el genoma del virus HCoV-229E (Thiel y col., 2001a). Para su generación se ha utilizado el sistema basado en el genoma de poxvirus como vector de clonaje y la posterior transcripción *in vitro* del RNA de coronavirus. Utilizando este sistema, los autores indican que, en ausencia del gen N, el replicón del HCoV-229E era funcional sólo a nivel de transcripción, mientras que en presencia del gen N se detectaba también la replicación del RNA, concluyendo que la nucleoproteína estaría implicada en replicación y no en transcripción (Schelle y col., 2005). Sin embargo, los resultados

aportados no proporcionan una evidencia suficiente que permita descartar la participación de la nucleoproteína en ambos procesos.

Recientemente, la proteína N ha sido descrita como una nueva chaperona de RNA. Estas proteínas favorecen el plegamiento correcto de RNAs y promueven la hibridación y desestabilización de dúplex de RNA. Esta observación ha llevado a proponer una nueva función de la proteína N durante la transcripción discontinua característica de coronavirus, facilitando el cambio de molde que ocurre durante la misma (Zúñiga y col., 2007). No obstante, la implicación de la proteína N de coronavirus en este proceso requiere de estudios adicionales. Los replicones del TGEV generados en este estudio, permitirán esclarecer a qué nivel, replicación o transcripción, actúa la proteína N, y caracterizar mediante mutagénesis puntual, dominios de la proteína implicados en la síntesis de RNA.

Además de la utilidad de los replicones derivados del TGEV para realizar estudios básicos de replicación, éstos constituyen una nueva generación de vectores de coronavirus basados en un genoma único, competente en replicación y deficiente en propagación, lo que favorece su aplicación biosegura. Estos vectores presentan las ventajas de una gran capacidad de clonaje, una expresión eficiente de genes heterólogos y la posibilidad de replicar en diferentes líneas celulares, incluyendo las células humanas 293T, donde el replicón no parece ser citopático. En contraste, los replicones mostraron un marcado efecto citopático en las células ST, BHK-pAPN y HeLa, lo que probablemente se debe a una mayor actividad del replicón en estas células respecto de las 293T. Sin embargo, no se puede descartar la existencia de factores específicos de hospedador o del tipo celular, que modulen la susceptibilidad de las diferentes células a la replicación del TGEV.

Los replicones derivados del TGEV pueden ser utilizados además para la evaluación de drogas y antivirales que interfieran con la replicación de coronavirus, tal como se ha mostrado con el replicón derivado del coronavirus humano HCoV-229E (Hertzog y col., 2004). El sistema de genética reversa desarrollado inicialmente para el coronavirus HCoV-229E, basado en la transfección inicial del RNA del replicón, presentaba una eficiencia considerablemente inferior a los replicones del TGEV generados en este estudio (Thiel y col., 2001a). Sin embargo, en trabajos posteriores se ha seleccionado un replicón no citopático del HCoV-229E mediante introducción de un gen de resistencia en la secuencia del gen de la replicasa, que ha aumentando notablemente la eficiencia y

aplicaciones del sistema (Hertzig y col., 2004; Schelle y col., 2005). La generación de replicones estables del TGEV, mediante su pase sucesivo en presencia de un gen de selección, permitirá optimizar en el laboratorio, de una forma similar, el sistema de replicones desarrollado para el TGEV y para el virus SARS-CoV (Almazán y col., 2006).

3. INTERACCIÓN DE PROTEÍNAS VIRALES Y CELULARES CON LOS EXTREMOS 5' Y 3' DEL GENOMA DEL TGEV

Las interacciones que establecen los virus con la célula hospedadora determinan la especificidad de hospedador, tejido y tipo celular, así como la patología de la infección, y su caracterización es un objetivo prioritario en la virología. Debido a su limitada capacidad codificante, los virus RNA reclutan proteínas de la maquinaria de biosíntesis de la célula hospedadora en cada etapa de su ciclo infectivo. A pesar de ello, sólo se conoce un número reducido de proteínas implicadas en la replicación de los virus RNA, debido a la dificultad de establecer una relación causal directa entre la función de las proteínas y la síntesis del RNA viral y la limitada disponibilidad de sistemas de replicación *in vitro*. El primer paso hacia la comprensión de la función de estas proteínas en el ciclo infectivo del virus es su identificación.

Hasta la fecha no existe ningún estudio de proteínas celulares implicadas en la síntesis de RNA en coronavirus del grupo 1, que incluye patógenos animales y humanos, y los estudios realizados con los coronavirus del grupo 2 MHV y BCoV, son muy limitados. En estos coronavirus, se han obtenido datos funcionales sólo en el caso de las proteínas hnRNP A1 (Shi y col., 2000), PABP (Spagnolo y Hogue, 2000), PTB (Huang y Lai, 1999) y aconitasa mitocondrial (Nanda y Leibowitz, 2001), que correlacionan cambios en su expresión y la síntesis de RNA.

En esta tesis se han aislado e identificado 30 proteínas celulares y virales por su interacción con los extremos del genoma del coronavirus TGEV, donde se localizan la mayor parte de los motivos de RNA esenciales para su replicación. Entre estas proteínas, la PABP, la hnRNP A1 y la hnRNP A2/B1 mostraron una unión preferencial por el extremo 3' del genoma, desde donde comienza la replicación con la síntesis de la cadena negativa. No obstante, las regiones reguladoras de los extremos 5' y 3' contienen, además de las señales *in cis* necesarias para la replicación, otros elementos

solapantes que regulan la transcripción, traducción, y estabilidad del RNA, por lo que las proteínas aisladas en este estudio constituyen candidatos cuya posible función en el ciclo infectivo de coronavirus ha de ser analizada en etapas posteriores.

Muchas de las proteínas de interacción con secuencias virales de interés han sido visualizadas tradicionalmente en experimentos de retardo en gel, aunque no se ha podido confirmar su identidad debido a la limitación de la cantidad y calidad de muestra que es necesaria para su identificación mediante proteómica. En el caso del coronavirus BCoV, se han detectado varias proteínas que interaccionan con la región 3'UTR, en la que mapean importantes señales de replicación, de las que sólo se conoce su masa molecular (Spagnolo y Hogue, 2000). El límite de detección de proteínas procedentes de bandas de gel mediante el uso de espectrometría de masas LC/MS/MS o MALDI-TOF está en el orden femtomolar, o en el rango de masa de 1-5 ng, similar al límite de detección de una tinción de un gel de proteínas con plata. Sin embargo, teniendo en cuenta la masa molecular de la proteína, la migración frecuente de múltiples proteínas en una misma banda de un gel monodimensional y la pérdida de material durante la digestión trípica, la cantidad de proteína necesaria para garantizar su identificación es cercana al límite de sensibilidad de la tinción con azul de Coomassie, es decir, entre 10-20 ng.

En este trabajo se ha puesto a punto una cromatografía para el aislamiento de proteínas por afinidad a RNA con un alto rendimiento, que ha permitido la identificación de 30 proteínas que interaccionan con los extremos del coronavirus TGEV.

Los RNAs de los extremos 5' y 3' del genoma que se utilizaron como ligando contienen la mayoría de los dominios esenciales para la replicación de minigenomas derivados del TGEV, con objeto de preservar estructuras de RNA de orden superior funcionalmente implicadas en la replicación, que se encuentran conservadas entre distintos coronavirus. Estos RNAs se marcaron mediante incorporación del análogo de ribonucleótido biotin-14-CTP durante la transcripción *in vitro*, en una relación 1:5.25 con respecto al nucleótido no marcado. Teniendo en cuenta la proporción de citosinas en las secuencias utilizadas (del 15-20%) y la menor eficiencia de incorporación del nucleótido modificado, estas condiciones permitieron el marcaje espaciado del RNA minimizando la posible interferencia con su estructura.

Un paso clave para la purificación específica de las proteínas de unión a RNA fue la eliminación de las proteínas de interacción inespecífica con la matriz de estreptavidina-

sefarosa, que se consiguió mediante el preaclaro previo de los extractos celulares en presencia de la resina. Sin embargo, no se puede descartar la eliminación de alguna proteína de interés en este paso. Estas proteínas necesitarían técnicas cromatográficas complementarias para su identificación. Otras posibles limitaciones de la metodología empleada son la utilización de extractos celulares citoplásmicos, que puede dar lugar a interacciones no fisiológicas debido a la descompartimentalización de la célula. Asimismo, la abundancia relativa de las proteínas en la célula, podría dificultar la identificación de factores celulares poco abundantes.

Como fuente de proteína se utilizaron extractos de citoplasma de células pertenecientes a tres especies distintas que soportan la replicación de coronavirus. El uso de extractos de células humanas HeLa permitió la identificación de un 99% de las bandas de proteínas visualizadas por tinción con Coomassie. Sin embargo, estas células no son susceptibles a la infección por el TGEV, aunque permiten un cierto grado de replicación viral, cuando los replicones derivados del TGEV se introdujeron por transfección en estas células. El proteoma de las células infectadas presenta diferencias en composición e interacciones intermoleculares, por lo que se decidió utilizar extractos de células infectadas con objeto de identificar las proteínas presentes en el contexto natural de la infección. Como limitación de esta aproximación, existe la posibilidad de que los RNAs virales secuestren los factores de interés a modo de competidores internos del sistema, aunque esta eliminación probablemente no sería completa y la utilización de un exceso molar de RNA ligando podría desplazar favorablemente estas interacciones.

Para la obtención de extractos de células infectadas se utilizaron las células porcinas ST, donde el virus replica con alta eficiencia, y las células de hámster BHK-pAPN, que soportan niveles altos de amplificación de los replicones del TGEV. La utilización de extractos procedentes de células ST o BHK-pAPN permitió la identificación del 50% y 45% de las proteínas respectivamente, una proporción muy inferior a la conseguida con las células humanas, debido a un menor número de secuencias del genoma de estas especies en las bases de datos. Recientemente, el consorcio para la secuenciación del genoma del cerdo doméstico (*Sus scrofa*) ha conseguido una cobertura de secuencia del 66%, aún sin ensamblar (Wernersson y col., 2005). En consecuencia, la identificación de proteínas de esta especie están limitadas porque las bases de datos no son completas y por el grado de homología de las secuencias de proteínas próximas con las de otras especies. En el caso de las células de hámster, cuyo genoma no ha sido secuenciado, las

identificaciones proceden por homología con especies como la humana, o la murina (*Mus musculus*), cuyo genoma está completo (Waterston y col., 2002), o la rata (*Rattus Norvegicus*) para la que se ha ensamblado una primera secuencia del genoma con una cobertura del 90% (Gibbs y col., 2004).

Una proporción elevada de las proteínas identificadas se unían indistintamente a los dos extremos del RNA del genoma y también a RNAs control del mismo tamaño pero con secuencia no relacionada con la del virus (datos no mostrados). Frecuentemente, las proteínas que interaccionan con RNA muestran cierta inespecificidad y se unen a motivos de pequeño tamaño que se encuentran abundantemente representados en las secuencias de RNA. Sin embargo, la especificidad en las interacciones RNA-proteína es clave para determinar su función biológica. La interacción específica *in vivo* puede estar favorecida por la localización subcelular, por su expresión diferencial, por la interacción con otros factores específicos o por la unión combinatoria y dinámica de distintas proteínas a cada RNA (Singh y Valcarcel, 2005). Debido a la especial naturaleza de estas interacciones, se decidió analizar la identidad de todas las proteínas que se unían a los RNAs de los extremos del virus, con el objetivo de identificar nuevas proteínas candidato cuya especificidad e implicación en el ciclo infectivo del virus deberá ser analizada posteriormente en ensayos funcionales.

Interesantemente, en algunos casos se pudieron detectar proteínas de unión preferencial a uno de los extremos del RNA y algunas de estas interacciones se observaron en los extractos celulares de las tres especies, lo que aportaba una especial significación a estas identificaciones. La proteína PABP y distintas isoformas de hnRNPs como la hnRNPA1 y la hnRNPA2/B1, mostraron una unión preferencial por el extremo 3' del genoma y, en el caso de la hnRNP A1, esta interacción diferencial se observó reproduciblemente en las células de las tres especies.

La proteína PABP está implicada en la traducción y en la regulación de la estabilidad de los mensajeros celulares. La unión de esta proteína al extremo 3' de los mRNAs media su circularización por interacción con factores de traducción que interaccionan con el extremo 5', lo que actúa sinérgicamente para favorecer su traducción eficiente (Jacobson y Peltz, 1996; Sachs y Varani, 2000). La PABP interacciona con el extremo 3' poliadenilado de diferentes virus RNA, jugando un papel en su traducción, y potencialmente también en su replicación. En el caso de poliovirus, la unión de la PABP a la cola de poliA estimula la traducción del genoma (Svitkin y col., 2001). Asimismo,

la asociación de la PABP con una estructura de RNA del extremo 5' y sus factores asociados, establece un puente 5'-3' que promueve la replicación del RNA (Herold y Andino, 2001). La PABP también interacciona con el extremo 3' de coronavirus del grupo II (BCV y MHV), y se ha descrito el requerimiento de la cola de poliA para la replicación de RNAs defectivos derivados (Spagnolo y Hogue, 2000). Los coronavirus podrían utilizar la PABP para estimular la traducción dependiente de cap de sus RNAs, y para regular el paso a la síntesis de RNA, por diferentes interacciones RNA-proteína y proteína-proteína, como ocurre en el caso de poliovirus, estableciendo una comunicación entre los extremos del genoma.

La proteína hnRNPA1 está implicada en la regulación del procesamiento alternativo de los mensajeros celulares y juega un papel en la síntesis de RNA en el coronavirus MHV, del grupo 2, (Furuya y Lai, 1993; Li y col., 1997; Shi y col., 2000; Zhang y Lai, 1995). Recientemente se ha descrito que en células con niveles indetectables de la proteína hnRNP A1, ésta puede ser reemplazada por otras isoformas tales como la hnRNPA2/B1 o la hnRNP A3 (Shi y col., 2003), lo que resuelve la controversia existente sobre el requerimiento de esta proteína para la replicación de MHV (Shen y Masters, 2001). En los resultados de esta tesis se han identificado las tres isoformas anteriores de hnRNPs por su interacción con el extremo 3' del RNA del coronavirus TGEV, del grupo 1, mostrando una mayor especificidad en el caso de la hnRNPA1 y hnRNPA2/B1. Estos resultados indican que estas interacciones podrían estar conservadas y tener relevancia funcional en coronavirus de los distintos grupos.

El resto de las proteínas se identificaron por su interacción con ambos extremos del genoma, presentaban un amplio rango de funciones relacionadas con la síntesis y el metabolismo del RNA celular y en muchos casos han sido asociadas a la replicación de otros virus RNA, incluidos los coronavirus.

Un grupo mayoritario de las proteínas identificadas están implicadas en procesamiento del pre-mRNA, incluyendo varias hnRNPs, como la hnRNP Q y hnRNP M y distintos componentes de los complejos de procesamiento del RNA, como las proteínas Prp4, Prp8, Prp3, el factor SF3B, o la helicasa DHX15. Paralelamente al desarrollo de este trabajo, la proteína hnRNP Q o SYNCRIP (del inglés, proteína citoplásmica de unión a RNA que interacciona con las sinaptotagminas), se ha descrito que interacciona con la 5'-UTR de polaridad positiva y negativa del genoma del coronavirus MHV (Choi y col., 2004). En este estudio, la expresión de un mutante C-

terminal de esta proteína produjo un efecto dominante negativo sobre la replicación de MHV y además, la reducción de su expresión utilizando RNAs interferentes produjo un retraso en la formación de sincitios y en la síntesis de proteínas y de RNA del virus, lo que apoya su posible implicación en el ciclo viral de MHV.

El significado biológico de la identificación de factores de *splicing* asociados a la replicación de un virus citoplásmico puede tener su base en el mecanismo de transcripción discontinua característico de los coronavirus. La existencia de RNAs subgenómicos de coronavirus 5' y 3' coterminales con el genoma, planteó inicialmente la posibilidad de un mecanismo de *splicing* para su generación, aunque las evidencias experimentales apoyan al modelo actual de transcripción discontinua, que no requiere *splicing* de la cadena positiva o negativa (Sawicki y col., 2007). No obstante, este modelo es perfectamente compatible con el requerimiento de factores en *trans* que aproximen en el espacio, de forma dinámica, las secuencias que regulan la síntesis de las distintas especies de RNA que, de otro modo, estarían demasiado distales debido al tamaño extraordinariamente grande de los genomas de coronavirus.

Otro grupo de proteínas identificadas está relacionado con el citoesqueleto celular. En este conjunto se incluyen la vimentina, que constituye filamentos intermedios, o la miosina, que posee actividad motora. La anexina A2, inicialmente descrita como una proteína de membrana regulada por calcio e implicada en procesos de exocitosis, se ha caracterizado recientemente como una nueva proteína de unión a RNA (Filipenko y col., 2004) implicada en el ensamblaje del virus HIV, mediante interacción con la proteína viral Gag (Ryzhova y col., 2006). La importancia de proteínas del citoesqueleto, especialmente las implicadas en el tráfico vesicular, es crucial a lo largo de múltiples pasos durante el ciclo de replicación de los virus, regulando el transporte intracelular del RNA, así como el ensamblaje, maduración y liberación de la partícula viral (Dohner y Sodeik, 2005). Las proteínas identificadas en este estudio, aisladas por afinidad a los extremos del genoma de coronavirus, podrían estar implicadas en el transporte del RNA entrante infeccioso hasta los complejos de replicación, y de los RNAs progenie para su traducción o empaquetamiento.

Algunas de las proteínas identificadas están implicadas en la traducción celular, incluyendo componentes ribosomales como la proteína S18, o enzimas del metabolismo del tRNA como la glutamil-prolil-tRNA sintetasa y podrían jugar un papel en la traducción preferente de los RNAs virales frente a los celulares.

Otras proteínas de unión a los extremos del genoma del TGEV son las chaperonas moleculares de inducción por choque térmico HSP70 mitocondrial (mtHSP70) y HSP90. Estos factores están implicados en el plegamiento de proteínas y se inducen en condiciones de estrés celular, como es el caso de las infecciones virales, donde desempeñan múltiples funciones, participando en procesos de síntesis de RNA, plegamiento de proteínas, y ensamblaje y maduración de los viriones (Mayer, 2005). Recientemente, la proteína mtHSP70 ha sido identificada junto con otras chaperonas, HSP60 y HSP40, por su interacción con el RNA del coronavirus MHV (Nanda y col., 2004), aunque no se ha asignado aún un papel funcional a esta interacción. En estudios posteriores, la proteína mtHSP70 ha sido implicada funcionalmente en la replicación de Tombusvirus en levaduras (Serva y Nagy, 2006) y la HSP90 estimula la replicación del virus Flock House en células de *Drosophila* (Kampmueller y Miller, 2005), por lo que podrían constituir factores estratégicamente conservados en la replicación de distintos virus RNA, incluidos los coronavirus.

Un grupo reducido de las proteínas identificadas están poco caracterizadas, como es el caso del antígeno RCD8. Este factor, inicialmente descrito como un autoantígeno humano asociado al síndrome de Sjogren, se ha caracterizado recientemente como un componente del complejo de eliminación de cap de los mensajeros celulares (Fenger-Gron y col., 2005). El antígeno RCD8 está considerado como una proteína que carece de motivos de unión a RNA, aunque la utilización de nuevas herramientas informáticas (Terribilini y col., 2006) permiten identificar una región característica de interfase proteína-RNA, comprendida entre los aminoácidos Thr⁸⁹⁷-Lys⁹²⁷ (datos no mostrados). Independientemente de un posible papel de esta proteína en la replicación de coronavirus, su purificación por afinidad a los extremos del genoma apoyaría que se tratara de una nueva proteína de unión a RNA.

Asimismo, también se identificaron varios factores de transcripción asociados a la RNA pol II celular, como son el factor de unión a la caja Y o la RNA helicasa A, o relacionados con la RNA pol III, como la ribonucleoproteína La o la nucleolina, todos ellos implicados en la replicación de diferentes virus RNA (Domitrovich y col., 2005; Holm y col., 2002; Huang y col., 2001; Isken y col., 2003). La replicación o transcripción de los virus RNA guarda cierto paralelismo con la transcripción del RNA celular dependiente de DNA (Lai, 1998). De forma similar a las polimerasas celulares, compuestas de múltiples subunidades, las polimerasas virales dependientes de RNA necesitan factores asociados que les confieran especificidad de molde, papel que

podrían desempeñar los factores de transcripción celulares asociados a la síntesis de RNA viral.

Aunque los coronavirus replican en células enucleadas, una fracción de la proteína N ha sido detectada en el nucleolo de células infectadas por los coronavirus TGEV, MHV e IBV, y en el caso del virus IBV, se ha observado la interacción de la nucleoproteína N con la nucleolina (Chen y col., 2002), por lo que esta proteína nuclear podría contribuir directa o indirectamente en la replicación de coronavirus. La proteína N se ha implicado en la síntesis del RNA viral y se ha aislado en este estudio por afinidad a los extremos del genoma del virus TGEV. Aunque la proteína de la nucleocápsida de coronavirus une diferentes RNAs de forma inespecífica, también presenta un componente de especificidad por secuencias de RNA virales, incluyendo regiones de ambos extremos del genoma de diferentes coronavirus (Nelson y Stohlman, 1993; Zhou y col., 1996).

Además de las proteínas celulares y la nucleoproteína, se detectaron varios péptidos tripticos en extractos de células ST infectadas que correspondían a la polimerasa viral dependiente de RNA, y que fueron obtenidos a partir de la digestión de una banda del tamaño correspondiente (100 kDa) (datos no mostrados). Sin embargo, esta identificación se obtuvo con un grado de significancia bajo, probablemente por tener una cobertura de secuencia estadísticamente insuficiente, debido al gran tamaño de las poliproteínas de la replicasa. Asimismo, la cantidad relativa de esta enzima en las células infectadas es probablemente baja, porque sólo se expresa cuando ocurre el salto de fase del ribosoma del gen de la replicasa, que da lugar a la poliproteína pp1ab. Para obtener una cantidad de proteína suficiente que facilite identificar las enzimas virales de síntesis de RNA será necesario enriquecer los extractos celulares en los complejos de replicación.

Los factores de transcripción RNA helicasa A y el factor de unión al *enhancer* de la interleukina 2 o NF45, pertenecientes al grupo de las proteínas NFAR, han sido aislados anteriormente por afinidad con secuencias de los extremos del genoma del virus BVDV y se han identificado como funcionalmente relevantes en la replicación viral. Ambas proteínas, parcialmente purificadas, median una comunicación entre los extremos 5' y 3' del RNA de BVDV, en un experimento de coprecipitación *in vitro* de RNA, por lo que se han propuesto como mediadores de la interacción entre los extremos del genoma (Isken y col., 2003). La purificación de las proteínas NFAR y de la PABP, por afinidad a los extremos del genoma de coronavirus TGEV, sugerían la posible implicación de

estos factores en una comunicación 5'-3' entre los extremos del genoma del TGEV, de forma similar a lo descrito para pestivirus y poliovirus.

La circularización del genoma de virus RNA mediada por proteínas se ha propuesto recientemente para pestivirus (Isken y col., 2003) y se ha caracterizado a nivel ultraestructural en el caso de flavivirus, donde de han descrito interacciones RNA-RNA a larga distancia (Álvarez y col., 2005). La circularización de genomas virales parece un mecanismo conservado en un número creciente de virus RNA que favorece la traducción y replicación de sus genomas (Edgil y Harris, 2006). Los extremos del genoma del TGEV no presentan complementariedad de secuencia. No obstante, en una predicción del plegamiento del genoma completo, la estructura termodinámicamente más estable muestra la proximidad espacial de los extremos 5' y 3' del RNA (Sgro y Palmenberg, comunicación personal). Esta conformación podría ser estabilizada *in vivo* mediante diferentes interacciones RNA-proteína y proteína-proteína.

En este estudio se analizó la posible interacción *in vitro* entre los extremos 5' y 3' del genoma del TGEV, utilizando un experimento de coprecipitación de RNA. Con esta aproximación se observó que la interacción entre los extremos 5' y 3' del genoma estaba mediada por proteínas, y que las proteínas celulares eran suficientes, al menos *in vitro*, para mediar esta interacción. Para estudiar la especificidad de las interacciones RNA-proteína implicadas en la comunicación 5'-3', la formación de los complejos se compitió con un exceso de RNA, homólogo o no relacionado. Se observó un desplazamiento proporcional de la señal de coprecipitación sólo en el primer caso, indicando que la interacción tiene un componente que es específico de RNA. Sin embargo, la utilización de distintas especies de RNA como competidor dio lugar a resultados variables (resultados no mostrados), desde la ausencia total de competición hasta niveles de competición próximos a lo observado con el RNA específico, lo que puede depender de la estructura de los diferentes RNAs.

Se ha identificado un elevado número de proteínas por su interacción con los extremos del genoma del TGEV en este trabajo. Estas proteínas son en principio candidatos responsables de la interacción entre los extremos del virus TGEV. En un estudio paralelo realizado con el virus MHV, se ha propuesto una posible interacción entre los extremos del genoma mediada por las proteínas hnRNP A1 y PTB, que se unen a ambos extremos del genoma, de polaridades opuestas, y además interaccionan entre sí (Huang y Lai, 2001). Adicionalmente, se ha observado la formación de un

complejo *in vitro* entre estas dos proteínas recombinantes y los dos extremos del genoma de MHV, aunque estos resultados no han sido confirmados con las proteínas endógenas. La proteína hnRNP A1, identificada en este estudio por su interacción preferencial con el extremo 3' del genoma del TGEV, podría estar implicada en mediar la interacción específica de los extremos de este virus, aunque la proteína PTB, que no se encuentra entre las proteínas identificadas, podría ser sustituida en el caso del virus TGEV por otro factor celular. En trabajos en curso durante la escritura de esta tesis se están identificando las proteínas que se encuentran en las fracciones de membrana donde tiene lugar la replicación viral de entre las proteínas identificadas. Para clasificar si estas proteínas tienen un papel relevante en la replicación de coronavirus se está inhibiendo su expresión utilizando RNAs interferentes. En este sistema se estudia el efecto de las distintas proteínas en la replicación viral analizando la amplificación de los replicones derivados del TGEV.

La circularización o interacción entre los extremos de genomas RNA, podría ser un mecanismo biológico que garantizara la traducción o replicación exclusiva a partir de genomas virales completos, para proteger los extremos del RNA de la acción de nucleasas, o para reciclar los factores de síntesis en sucesivas rondas de replicación o traducción. Teniendo en cuenta los datos disponibles hasta la fecha acerca de la replicación de coronavirus, incluyendo los obtenidos en este trabajo, se podría postular un modelo de trabajo para la replicación del genoma de coronavirus (Fig. 42) que es compatible con las siguientes observaciones experimentales: a) la replicación ocurre asociada a membranas intracelulares; b) las señales de replicación en *cis* se localizan en los extremos 5' y 3' del genoma; c) el extremo 3' parece ser suficiente para la síntesis de la cadena negativa y ambos extremos participan durante para la síntesis de la cadena positiva; d) la síntesis de la cadena positiva es asimétrica a partir de intermediarios de replicación de polaridad negativa; e) la proteína hnRNP A1 interacciona específicamente con el extremo 3', de los virus MHV y TGEV, y puede jugar un papel en la síntesis del RNA; f) la proteína PABP interacciona con la cola de poliA; g) pueden existir interacciones adicionales RNA-proteína y proteína proteína, participando proteínas como la PTB, la HSP70 o la hnRNP Q; h) los extremos del genoma probablemente interaccionan durante la replicación; i) la proteína N juega un papel esencial en la síntesis del RNA.

El modelo de trabajo para la replicación de coronavirus que se propone (Fig. 42) comenzaría con la síntesis de la cadena negativa desde el extremo 3' del genoma,

proceso en el que participarían las señales en *cis* localizadas en este extremo del genoma molde (Fig. 42A). Las proteínas específicamente asociadas al extremo 3', como la PABP o la hnRNP A1 podrían contribuir al reconocimiento de este extremo por la RpRd viral. La helicasa del virus actuaría por delante de la polimerasa, deshaciendo estructuras en el RNA. A medida que avanza la síntesis de la cadena negativa, las dos cadenas permanecerían asociadas formando intermedios replicativos. A su vez las proteínas que se unen a los extremos del genoma podrían interaccionar entre sí, directa

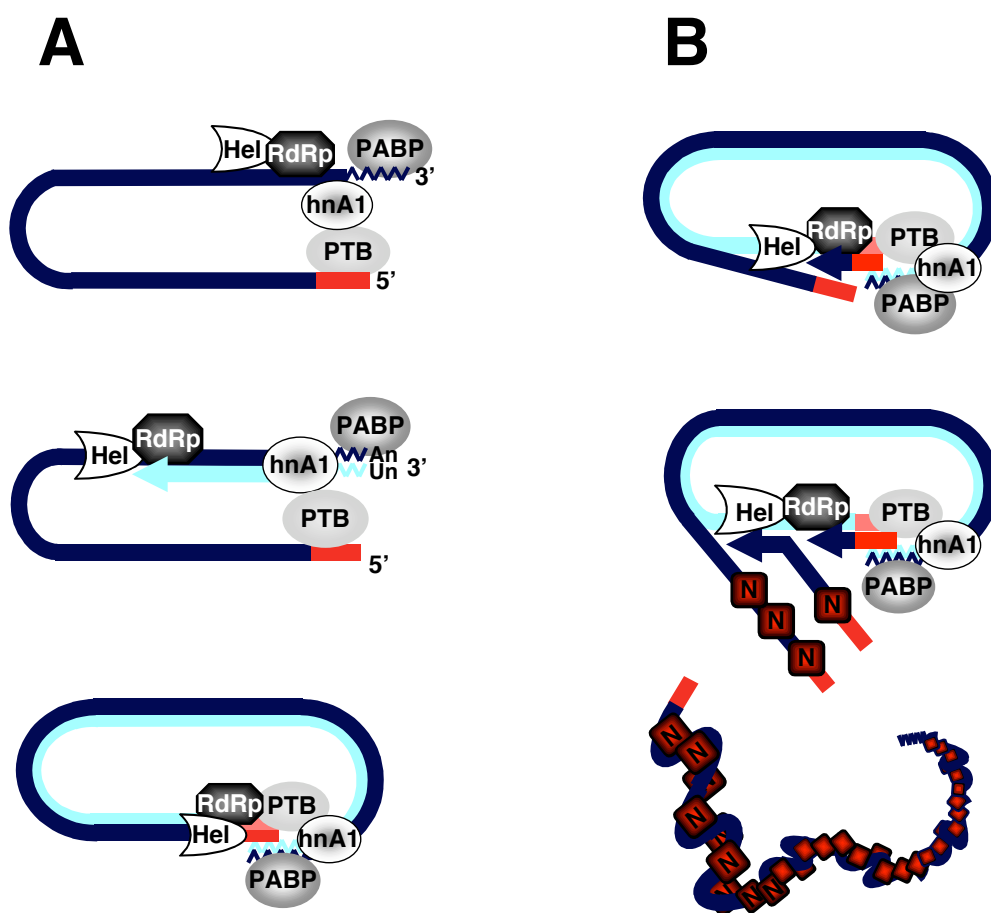


Figura 42. Modelo de trabajo de la replicación de coronavirus.(A) Síntesis de la cadena negativa. Se representa el genoma por una línea azul oscura con la secuencia líder en rojo. La cadena negativa se representa por una línea azul claro. Se muestran las proteínas específicamente aisladas por afinidad al extremo 3' del TGEV en este trabajo y la PTB, que interacciona con el genoma de MHV, aunque en el caso del TGEV no se ha identificado y podría estar reemplazada por otro factor. Hel, helicasa; hnA1, hnRNP A1; An y Un, poliA y poliU, respectivamente (B) Síntesis de la cadena positiva.

o indirectamente, lo que situaría a los extremos 5' y 3' espacialmente próximos. Esta circularización del genoma podría estar mediada en el coronavirus MHV por la interacción de las proteínas PTB y hnRNP A1. La conformación circular permitiría el reciclaje de los factores de síntesis, lo que aumentaría la eficiencia de las re-

iniciaciones. Esto favorecería el uso de esta estructura como molde eficiente de replicación, dando lugar a la síntesis asimétrica de múltiples cadenas positivas (Fig. 42B). La síntesis de las cadenas positivas podría requerir un desplazamiento de molde, asistido por la actividad de la helicasa para abrir el intermedio de replicación. En esta etapa intervendrían las secuencias en *cis* de ambos extremos del genoma y presumiblemente, las proteínas asociadas. Además del posible papel de la proteína N como parte del complejo de replicación durante la síntesis del RNA, ésta proteína podría facilitar la liberación de los genomas nacientes, a los que se asociaría para su protección y posterior empaquetamiento.

Para comprobar la corrección de este modelo de trabajo será conveniente conocer las proteínas, tanto virales como celulares, que forman parte del complejo de replicación y la obtención de datos funcionales de las mismas.

VIII. CONCLUSIONES

1. Un cambio puntual de aminoácido en la posición 108 de las poliproteínas de la replicasa, localizado en el sitio de procesamiento nsp1/nsp2, afecta de forma diferencial a la replicación del coronavirus TGEV y a la amplificación de minigenomas derivados del mismo. La replicación eficiente del TGEV requería de aminoácidos que permitían el procesamiento nsp1/nsp2, a diferencia de la amplificación de los minigenomas, que requería mutaciones que afectaban al sitio de procesamiento, lo que sugiere diferentes requerimientos para la replicación en *cis* y en *trans*.

2. La comparación de distintos métodos para la transfección de células eucarióticas con DNAs de gran tamaño (BACs) ha permitido seleccionar un método de lipofección con el que se ha conseguido un aumento de cinco unidades logarítmicas en el rescate de virus TGEV en el pase P0 a partir de la transfección del clon infectivo, alcanzándose títulos de hasta 10^8 ufp/ml. La optimización de la transfección de BACs ha facilitado el análisis de la actividad de los replicones derivados del TGEV.

3. La síntesis del RNA del TGEV requiere el gen de la replicasa, los extremos 5' y 3' del genoma y la proteína N. La proteína N es esencial para la actividad eficiente de replicones derivados del TGEV y puede ser aportada tanto en *cis* como en *trans*. Su actividad depende de la proteína en sí misma y no del mRNA que la codifica.

4. Se han desarrollado vectores virales eficientes en replicación y deficientes en propagación basados en un genoma único. Estos replicones fueron funcionales en células de diferentes especies, incluyendo la humana, en las que mostraron un reducido efecto citopático y altos niveles de expresión de la proteína heteróloga EGFP. Los replicones generados constituyen vectores virales bioseguros, útiles en estudios básicos de replicación y para la evaluación de antivirales.

5. El estudio de la interacción de los extremos 5' y 3' del genoma del TGEV con proteínas virales y celulares han permitido identificar 30 proteínas, de las cuales la hnRNP A1, hnRNP A2/B1 y PABP mostraron una unión preferencial al extremo 3' del genoma. Entre las proteínas identificadas se incluyen factores esenciales en la replicación del TGEV, como la nucleoproteína N y la RNA polimerasa viral, y las proteínas celulares hnRNP A1, hnRNP Q y PABP que han sido implicadas en la

Conclusiones

replicación de otros coronavirus. La mayoría de las restantes proteínas identificadas están implicadas en la síntesis y procesamiento del RNA celular y son candidatos nuevos para el estudio de las interacciones coronavirus-huésped.

6. Proteínas celulares aisladas por su capacidad de unión a los extremos del genoma del TGEV son suficientes para mediar la interacción de los mismos *in vitro*. Se ha postulado que la replicación del genoma de coronavirus implica su circularización.

IX. BIBLIOGRAFÍA

- Agapov, E. V., Frolov, I., Lindenbach, B. D., Pragai, B. M., Schlesinger, S. and Rice, C. M.** 1998. Noncytopathic Sindbis virus RNA vectors for heterologous gene expression. *Proc. Natl. Acad. Sci. USA* **95**, 12989-12994.
- Ahlquist, P., Noueiry, A. O., Lee, W. M., Kushner, D. B. and Dye, B. T.** 2003. Host factors in positive-strand RNA virus genome replication. *J. Virol.* **77**, 8181-6.
- Almazán, F., DeDiego, M. L., Galán, C., Escors, D., Alvarez, E., Ortego, J., Sola, I., Zúñiga, S., Alonso, S., Moreno, J. L. et al.** 2006. Construction of a SARS-CoV infectious cDNA clone and a replicon to study coronavirus RNA synthesis. *J. Virol.* **80**, 10900-10906.
- Almazán, F., González, J. M., Péñzes, Z., Izeta, A., Calvo, E., Plana-Durán, J. and Enjuanes, L.** 2000. Engineering the largest RNA virus genome as an infectious bacterial artificial chromosome. *Proc. Natl. Acad. Sci. USA* **97**, 5516-5521.
- Alonso, S., Sola, I., Teifke, J., Reimann, I., Izeta, A., Balach, M., Plana-Durán, J., Moormann, R. J. M. and Enjuanes, L.** 2002. In vitro and in vivo expression of foreign genes by transmissible gastroenteritis coronavirus-derived minigenomes. *J. Gen. Virol.* **83**, 567-579.
- Álvarez, D. E., Lodeiro, M. F., Luduena, S. J., Pietrasanta, L. I. and Gamarnik, A. V.** 2005. Long-range RNA-RNA interactions circularize the dengue virus genome. *J. Virol.* **79**, 6631-6643.
- Andino, R., Silvera, D., Suggett, S. D., Achacoso, P. L., Miller, C. J., Baltimore, D. and Feinberg, M. B.** 1994. Engineering poliovirus as a vaccine vector for the expression of diverse antigens. *Science* **265**, 1448-1451.
- Ballesteros, M. L., Sanchez, C. M. and Enjuanes, L.** 1997. Two amino acid changes at the N-terminus of transmissible gastroenteritis coronavirus spike protein result in the loss of enteric tropism. *Virology* **227**, 378-388.
- Baranov, P. V., Henderson, C. M., Anderson, C. B., Gesteland, R. F., Atkins, J. F. and Howard, M. T.** 2005. Programmed ribosomal frameshifting in decoding the SARS-CoV genome. *Virology* **332**, 498-510.
- Baric, R. S., Nelson, G. W., Fleming, J. O., Deans, R. J., Keck, J. G., Casteel, N. and Stohlman, S. A.** 1988. Interactions between coronavirus nucleocapsid protein and viral RNAs: implications for viral transcription. *J. Virol.* **62**, 4280-4287.
- Ben-David, Y., Bani, M. R., Chabot, B., De Koven, A. and Bernstein, A.** 1992. Retroviral insertions downstream of the heterogeneous nuclear ribonucleoprotein A1 gene in erythroleukemia cells: evidence that A1 is not essential for cell growth. *Mol. Cell. Biol.* **12**, 4449-55.
- Bergamini, G., Preiss, T. and Hentze, M. W.** 2000. Picornavirus IRESes and the poly(A) tail jointly promote cap-independent translation in a mammalian cell-free system. *RNA* **6**, 1781-1790.
- Bertling, W. M., Gareis, M., Paspaleeva, V., Zimmer, A., Kreuter, J., Nurnberg, E. and Harrer, P.** 1991. Use of liposomes, viral capsids, and nanoparticles as DNA carriers. *Biotechnol. Appl. Biochem.* **13**, 390-405.
- Blumenthal, T. and Carmichael, G. G.** 1979. RNA replication: function and structure of Qbeta-replicase. *Annu. Rev. Biochem.* **48**, 525-48.

- Boguszewska-Chachulska, A. M. and Haenni, A. L.** 2005. RNA viruses redirect host factors to better amplify their genome. *Adv. Virus Res.* **65**, 29-61.
- Bol, J. F.** 1999. Alfalfa mosaic virus and ilarviruses: involvement of coat protein in multiple steps of the replication cycle. *J Gen Virol* **80 (Pt 5)**, 1089-102.
- Bonilla, P. J., Hughes, S. A. and Weiss, S. R.** 1997. Characterization of a second cleavage site and demonstration of activity in trans by the papain-like proteinase of the murine coronavirus mouse hepatitis virus strain A59. *J. Virol.* **71**, 900-909.
- Boussif, O., Lezoualc'h, F., Zanta, M. A., Mergny, M. D., Scherman, D., Demeneix, B. and Behr, J. P.** 1995. A versatile vector for gene and oligonucleotide transfer into cells in culture and in vivo: polyethylenimine. *Proc. Natl. Acad. Sci. U S A* **92**, 7297-301.
- Brian, D. A. and Baric, R. S.** 2005. Coronavirus genome structure and replication. *Curr. Top. Microbiol. Immunol.* **287**, 1-30.
- Brierley, I., Digard, P. and Inglis, S. C.** 1989. Characterization of an efficient coronavirus ribosomal frameshifting signal: requirement for an RNA pseudoknot. *Cell* **57**, 537-547.
- Brockway, S. M. and Denison, M. R.** 2005. Mutagenesis of the murine hepatitis virus nsp1-coding region identifies residues important for protein processing, viral RNA synthesis, and viral replication. *Virology* **340**, 209-23.
- Brockway, S. M., Lu, X. T., Peters, T. R., Dermody, T. S. and Denison, M. R.** 2004. Intracellular localization and protein interactions of the gene 1 protein p28 during mouse hepatitis virus replication. *J. Virol.* **78**, 11551-62.
- Bronstein, I., Fortin, J. J., Voyta, J. C., Juo, R.-R., Edwards, B., Olenses, C. E. M., Lijam, N. and Kricka, L. J.** 1994. Chemiluminescent reporter gene assays: sensitive detection of the GUS and SEAP gene products. *BioTechniques* **17**, 172-177.
- Bronstein, I., Martín, C. S., Fortin, J. J., Olesen, C. E. M. and Voyta, J. C.** 1996. Chemiluminescence: sensitive detection technology for reporter gene assays. *Clin. Chem.* **42**, 1542-1546.
- Caley, I. J., Betts, M. R., Davis, N. L., Swanstrom, R., Frelinger, J. A. and Johnston, R. E.** 1999. Venezuelan equine encephalitis virus vectors expressing HIV-1 proteins: vector design strategies for improved vaccine efficacy. *Vaccine* **17**, 3124-3135.
- Calvo, E., Escors, D., Lopez, J. A., González, J. M., Álvarez, A., Arza, E. and Enjuanes, L.** 2005. Phosphorylation and subcellular localization of transmissible gastroenteritis virus nucleocapsid protein in infected cells. *J. Gen. Virol.* **86**, 2255-2267.
- Casais, R., Thiel, V., Siddell, S. G., Cavanagh, D. and Britton, P.** 2001. Reverse genetics system for the avian coronavirus infectious bronchitis virus. *J. Virol.* **75**, 12359-12369.

- Cavanagh, D., Brian, D. A., Britton, P., Enjuanes, L., Horzinek, M. C., Lai, M. M. C., Laude, H., Plagemann, P. G. W., Siddell, S., Spaan, W. et al.** 1997. *Nidovirales*: a new order comprising *Coronaviridae* and *Arteriviridae*. *Arch. Virol.* **142**, 629-635.
- Chambers, T. J., Nestorowicz, A., Mason, P. W. and Rice, C. M.** 1999. Yellow fever/Japanese encephalitis chimeric viruses: construction and biological properties. *J. Virol.* **73**, 3095-3101.
- Chang, C. K., Sue, S. C., Yu, T. H., Hsieh, C. M., Tsai, C. K., Chiang, Y. C., Lee, S. J., Hsiao, H. H., Wu, W. J., Chang, W. L. et al.** 2006. Modular organization of SARS coronavirus nucleocapsid protein. *J. Biomed. Sci.* **13**, 59-72.
- Chang, R. Y., Hofmann, M. A., Sethna, P. B. and Brian, D. A.** 1994. A *cis*-acting function for the coronavirus leader in defective interfering RNA replication. *J. Virol.* **68**, 8223-8231.
- Chang, R. Y., Krishnan, R. and Brian, D. A.** 1996. The UCUAAC promoter motif is not required for high-frequency leader recombination in bovine coronavirus defective interfering RNA. *J. Virol.* **70**, 2720-9.
- Chen, C. J. and Makino, S.** 2004. Murine coronavirus replication induces cell cycle arrest in G₀/G₁ phase. *J. Virol.* **78**, 5658-5669.
- Chen, C. J., Sugiyama, K., Kubo, H., Huang, C. and Makino, S.** 2004. Murine coronavirus nonstructural protein p28 arrests cell cycle in G₀/G₁ phase. *J. Virol.* **78**, 10410-10419.
- Chen, H., Gill, A., Dove, B. K., Emmett, S. R., Kemp, C. F., Ritchie, M. A., Dee, M. and Hiscox, J. A.** 2005. Mass spectroscopic characterization of the coronavirus infectious bronchitis virus nucleoprotein and elucidation of the role of phosphorylation in RNA binding by using surface plasmon resonance. *J. Virol.* **79**, 1164-1179.
- Chen, H., Wurm, T., Britton, P., Brooks, G. and Hiscox, J. A.** 2002. Interaction of the coronavirus nucleoprotein with nucleolar antigens and the host cell. *J. Virol.* **76**, 5233-5250.
- Choi, K. S., Mizutani, A. and Lai, M. M.** 2004. SYNCRIP, a member of the heterogeneous nuclear ribonucleoprotein family, is involved in mouse hepatitis virus RNA synthesis. *J. Virol.* **78**, 13153-13162.
- Coley, S. E., Lavi, E., Sawicki, S. G., Fu, L., Schelle, B., Karl, N., Siddell, S. G. and Thiel, V.** 2005. Recombinant mouse hepatitis virus strain a59 from cloned, full-length cDNA replicates to high titers in vitro and is fully pathogenic in vivo. *J. Virol.* **79**, 3097-3106.
- Cologna, R., Spagnolo, J. F. and Hogue, B. G.** 2000. Identification of nucleocapsid binding sites within coronavirus-defective genomes. *Virology* **277**, 235-249.
- Correa, I., Jiménez, G., Suñé, C., Bullido, M. J. and Enjuanes, E.** 1988. Antigenic structure of the E2 glycoprotein from transmissible gastroenteritis coronavirus. *Virus Res.* **10**, 77-94.
- Craig, A. W., Haghighat, A., Yu, A. T. and Sonenberg, N.** 1998. Interaction of polyadenylate-binding protein with the eIF4G homologue PAIP enhances translation. *Nature* **392**, 520-3.

- Dalton, K., Casais, R., Shaw, K., Stirrups, K., Evans, S., Britton, P., Brown, T. D. K. and Cavanagh, D.** 2001. *cis*-acting sequences required for coronavirus infectious bronchitis virus defective-RNA replication and packaging. *J. Virol.* **75**, 125-133.
- de Vries, A. A. F., Horzinek, M. C., Rottier, P. J. M. and de Groot, R. J.** 1997. The genome organization of the *Nidovirales*: similarities and differences between arteri-, toro-, and coronaviruses. *Semin. Virol.* **8**, 33-47.
- Delmas, B., Gelfi, J., L'Haridon, R., Vogel, L. K., Norén, O. and Laude, H.** 1992. Aminopeptidase N is a major receptor for the enteropathogenic coronavirus TGEV. *Nature* **357**, 417-420.
- Delmas, B., Gelfi, J. and Laude, H.** 1986. Antigenic structure of transmissible gastroenteritis virus. II. Domains in the peplomer glycoprotein. *J. Gen. Virol.* **67**, 1405-1418.
- Delmas, B. and Laude, H.** 1990. Assembly of coronavirus spike protein into trimers and its role in epitope expression. *J. Virol.* **64**, 5367-5375.
- Denison, M. R., Spaan, W. J. M., van der Meer, Y., Gibson, C. A., Sims, A. C., Prentice, E. and Lu, S. T.** 1999. The putative helica of the coronavirus mouse hepatitis virus is processed from the replicase gene polyprotein and localizes in complexes that are active in viral RNA synthesis. *J. Virol.* **73**, 6862-6871.
- Denison, M. R., Yount, B., Brockway, S. M., Graham, R. L., Sims, A. C., Lu, X. T. and Baric, R. S.** 2004. Cleavage between replicase proteins p28 and p65 of mouse hepatitis virus is not required for virus replication. *J. Virol.*, 5957-5965.
- Denison, M. R., Zoltick, P. W., Hughes, S. A., Giangreco, B., Olson, A. L., Perlman, S., Leibowitz, J. L. and Weiss, S. R.** 1992. Intracellular processing of the N-terminal ORF-1a proteins of the coronavirus MHV-A59 requires multiple proteolytic events. *Virology* **189**, 274-284.
- Dohner, K. and Sodeik, B.** 2005. The role of the cytoskeleton during viral infection. *Curr. Top. Microbiol. Immunol.* **285**, 67-108.
- Domitrovich, A. M., Diebel, K. W., Ali, N., Sarker, S. and Siddiqui, A.** 2005. Role of La autoantigen and polypyrimidine tract-binding protein in HCV replication. *Virology* **335**, 72-86.
- Dong, S. and Baker, S.** 1994. Determinants of the p28 cleavage site recognized by the first papain-like cysteine proteinase of murine coronavirus. *Virology* **204**, 541-549.
- Doyle, L. P. and Hutchings, L. M.** 1946. A transmissible gastroenteritis in pigs. *J. Amer. Vet. Med. Assoc.* **108**, 257-259.
- Drosten, C., Günther, S., Preiser, W., van der Werf, S., Brodt, H.-R., Becker, S., Rabenau, H., Panning, M., Kolesnikova, L., Fouchier, R. A. M. et al.** 2003. Identification of a novel coronavirus in patients with severe acute respiratory syndrome. *N. Engl. J. Med.* **348**, 1967-1976.
- Edgil, D. and Harris, E.** 2006. End-to-end communication in the modulation of translation by mammalian RNA viruses. *Virus. Res.* **119**, 43-51.
- Eleouet, J. F., Rasschaert, D., Lambert, P., Levy, L., Vende, P. and Laude, H.** 1995. Complete sequence (20 kilobases) of the polyprotein-encoding gene 1 of transmissible gastroenteritis virus. *Virology* **206**, 817-22.

- Enjuanes, L., Almazán, F. and Ortego, J.** 2003. Virus-based vectors for gene expression in mammalian cells: Coronavirus. In *Gene Transfer and Expression in Mammalian cells*, ed. S. C. Makrides, pp. 151-168: Elsevier Science B.V.
- Enjuanes, L., Almazán, F., Sola, I. and Zúñiga, S.** 2006. Biochemical aspects of coronavirus replication and virus-host interaction. *Annu. Rev. Microbiol.* **60**, 211-230.
- Enjuanes, L., Brian, D., Cavanagh, D., Holmes, K., Lai, M. M. C., Laude, H., Masters, P., Rottier, P., Siddell, S. G., Spaan, W. J. M. et al.** 2000a. Coronaviridae. San Diego, California: Academic Press.
- Enjuanes, L., Sola, I., Alonso, S., Escors, D. and Zúñiga, S.** 2005. Coronavirus reverse genetics and development of vectors for gene expression. *Curr. Top. Microbiol. Immunol.* **287**, 161-97.
- Enjuanes, L., Spaan, W., Snijder, E. and Cavanagh, D.** 2000b. *Nidovirales*. In *Virus taxonomy. Classification and nomenclature of viruses*, eds. M. H. V. van Regenmortel C. M. Fauquet D. H. L. Bishop E. B. Carsten M. K. Estes S. M. Lemon D. J. McGeoch J. Maniloff M. A. Mayo C. R. Pringle et al., pp. 827-834. San Diego, California: Academic Press.
- Enjuanes, L. and Van der Zeijst, B. A. M.** 1995. Molecular basis of transmissible gastroenteritis coronavirus epidemiology. In *The Coronaviridae*, (ed. S. G. Siddell), pp. 337-376. New York: Plenum Press.
- Escors, D., Camafeita, E., Ortego, J., Laude, H. and Enjuanes, L.** 2001. Organization of two transmissible gastroenteritis coronavirus membrane protein topologies within the virion and core. *J. Virol.* **75**, 12228-12240.
- Fan, H., Ooi, A., Tan, Y. W., Wang, S., Fang, S., Liu, D. X. and Lescar, J.** 2005. The nucleocapsid protein of coronavirus infectious bronchitis virus: crystal structure of its N-terminal domain and multimerization properties. *Structure* **13**, 1859-68.
- Felgner, P. L., Gadek, T. R., Holm, M., Roman, R., Chan, H. W., Wenz, M., Northrop, J. P., Ringold, G. M. and Danielsen, M.** 1987. Lipofection: a highly efficient, lipid-mediated DNA-transfection procedure. *Proc. Natl. Acad. Sci. U S A* **84**, 7413-7.
- Fenger-Gron, M., Fillman, C., Norrild, B. and Lykke-Andersen, J.** 2005. Multiple processing body factors and the ARE binding protein TTP activate mRNA decapping. *Mol. Cell. Biol.* **20**, 905-15.
- Filipenko, N. R., MacLeod, T. J., Yoon, C. S. and Waisman, D. M.** 2004. Annexin A2 is a novel RNA-binding protein. *J. Biol. Chem.* **279**, 8723-31.
- Fominaya, J., Gasset, M., Garcia, R., Roncal, F., Albar, J. P. and Bernad, A.** 2000. An optimized amphiphilic cationic peptide as an efficient non-viral gene delivery vector. *J. Gene. Med.* **2**, 455-64.
- Frolov, I., Agapov, E., Hoffman, T. A., Prágai, B. M., Lippa, M., Schlesinger, S. and Rice, C. M.** 1999. Selection of RNA replicons capable of persistent noncytopathic replication in mammalian cells. *J. Virol.* **73**, 3854-3865.

- Frolov, I., Hoffman, T. A., Prágai, B. M., Dryga, S. A., Huang, H. V., Schlesinger, S. and Rice, C. M.** 1996. Alphavirus-based expression vectors: Strategies and applications. *Proc. Natl. Acad. Sci. USA* **93**, 11371-11377.
- Furuya, T. and Lai, M. M. C.** 1993. Three different cellular proteins bind to complementary sites on the 5'-end-positive and 3'-end negative strands of mouse hepatitis virus RNA. *J. Virol.* **67**, 7215-7222.
- Gamarnik, A. V. and Andino, R.** 1997. Two functional complexes formed by KH domain containing proteins with the 5' noncoding region of poliovirus RNA. *RNA* **3**, 882-892.
- Gibbs, R. A., Weinstock, G. M., Metzker, M. L., Muzny, D. M., Sodergren, E. J., Scherer, S., Scott, G., Steffen, D., Worley, K. C., Burch, P. E. et al.** 2004. Genome sequence of the Brown Norway rat yields insights into mammalian evolution. *Nature* **428**, 493-521.
- Ginalski, K., Godzik, A. and Rychlewski, L.** 2006. Novel SARS Unique AdoMet-Dependent Methyltransferase. *Cell Cycle* **5**, 2414-2416.
- González, J. M., Gomez-Puertas, P., Cavanagh, D., Gorbalenya, A. E. and Enjuanes, L.** 2003. A comparative sequence analysis to revise the current taxonomy of the family Coronaviridae. *Arch. Virol.* **148**, 2207-2235.
- González, J. M., Penzes, Z., Almazan, F., Calvo, E. and Enjuanes, L.** 2002. Stabilization of a full-length infectious cDNA clone of transmissible gastroenteritis coronavirus by the insertion of an intron. *J. Virol.* **76**, 4655-4661.
- Gorbalenya, A. E., Enjuanes, L., Ziebuhr, J. and Snijder, E. J.** 2006. Nidovirales: evolving the largest RNA virus genome. *Virus Res.* **117**, 17-37.
- Gosert, R., Kanjanahaluethai, A., Egger, D., Bienz, K. and Baker, S. C.** 2002. RNA replication of mouse hepatitis virus takes place at double-membrane vesicles. *J. Virol.* **76**, 3697-3708.
- Graham, R. L. and Denison, M. R.** 2006. Replication of murine hepatitis virus is regulated by papain-like proteinase 1 processing of nonstructural proteins 1, 2, and 3. *J. Virol.* **80**, 11610-20.
- Graham, R. L., Sims, A. C., Brockway, S. M., Baric, R. S. and Denison, M. R.** 2005. The nsp2 replicase proteins of murine hepatitis virus and severe acute respiratory syndrome coronavirus are dispensable for viral replication. *J. Virol.* **79**, 13399-411.
- Han, Y. S., Chang, G. G., Juo, C. G., Lee, H. J., Yeh, S. H., Hsu, J. T. and Chen, X.** 2005. Papain-like protease 2 (PLP2) from severe acute respiratory syndrome coronavirus (SARS-CoV): expression, purification, characterization, and inhibition. *Biochemistry* **44**, 10349-59.
- Hanahan, D.** 1985. Techniques for transformation of *E. coli*. In *DNA Cloning. A Practical Approach*, vol. 1 (ed. D. H. Glover), pp. 109-136. Oxford: IRL Press Oxford.
- Herold, J. and Andino, R.** 2000. Poliovirus requires a precise 5' end for efficient positive-strand RNA synthesis. *J. Virol.* **74**, 6394-6400.
- Herold, J. and Andino, R.** 2001. Poliovirus RNA replication requires genome circularization through a protein-protein bridge. *Mol. Cell.* **7**, 581-591.

- Herold, J., Gorbalenya, A. E., Thiel, V., Schelle, B. and Siddell, S. G.** 1998. Proteolytic processing at the amino terminus of human coronavirus 229E gene 1-encoded polyproteins: identification of a papain-like proteinase and its substrate. *J. Virol.* **72**, 910-918.
- Hertzog, T., Sacandella, E., Schelle, B., Ziebuhr, J., Siddell, S., Ludewig, B. and Thiel, V.** 2004. Rapid identification coronavirus replicase inhibitors using a selectable replicon RNA. *J. Gen. Virol.* **85**, 1717-1725.
- Hiscox, J. A., Wurm, T., Wilson, L., Britton, P., Cavanagh, D. and Brooks, G.** 2001. The coronavirus infectious bronchitis virus nucleoprotein localizes to the nucleolus. *J. Virol.* **75**, 506-512.
- Holm, P. S., Bergmann, S., Jurchott, K., Lage, H., Brand, K., Ladhoff, A., Mantwill, K., Curiel, D. T., Dobbstein, M., Dietel, M. et al.** 2002. YB-1 relocates to the nucleus in adenovirus-infected cells and facilitates viral replication by inducing E2 gene expression through the E2 late promoter. *J. Biol. Chem.* **277**, 10427-34.
- Hsue, B. and Masters, P. S.** 1997. A bulged stem-loop structure in the 3' untranslated region of the genome of the coronavirus mouse hepatitis virus is essential for replication. *J. Virol.* **71**, 7567-7578.
- Huang, P. and Lai, M. M. C.** 1999. Polypyrimidine tract-binding protein binds to the complementary strand of the mouse hepatitis virus 3' untranslated region, thereby altering RNA conformation. *J. Virol.* **73**, 9110-9116.
- Huang, P. and Lai, M. M. C.** 2001. Heterogeneous nuclear ribonucleoprotein A1 binds to the 3'-untranslated region and mediates potential 5'-3'-end cross talks of mouse hepatitis virus RNA. *J. Virol.* **75**, 5009-5017.
- Huang, Q., Yu, L., M., P. A., Gunasekera, A., Liu, Z., Xu, N., Hajduk, P., Mack, J., W., F. S. and Olejniczak, E. T.** 2004. Structure of the N-terminal RNA-binding domain of the SARS CoV nucleocapsid protein. *Biochemistry* **20**, 6059-6063.
- Huang, W. H., Yung, B. Y., Syu, W. J. and Lee, Y. H.** 2001. The nucleolar phosphoprotein B23 interacts with hepatitis delta antigens and modulates the hepatitis delta virus RNA replication. *J. Biol. Chem.* **276**, 25166-75.
- Hughes, S. A., Bonilla, P. J. and Weiss, S. R.** 1995. Identification of the murine coronavirus p28 cleavage site. *J. Virol.* **69**, 809-13.
- Hurst, K. R., Kuo, L., Koetzner, C. A., Ye, R., Hsue, B. and Masters, P. S.** 2005. A major determinant for membrane protein interaction localizes to the carboxy-terminal domain of the mouse coronavirus nucleocapsid protein. *J. Virol.* **79**, 13285-13297.
- Iborra, F. J., Escargueil, A. E., Kwek, K. Y., Akoulitchiev, A. and Cook, P. R.** 2004. Molecular cross-talk between the transcription, translation, and nonsense-mediated decay machineries. *J. Cell. Sci.* **117**, 899-906.
- Imbert, I., Guillemot, J. C., Bourhis, J. M., Bussetta, C., Coutard, B., Egloff, M. P., Ferron, F., Gorbalenya, A. E. and Canard, B.** 2006. A second, non-canonical RNA-dependent RNA polymerase in SARS Coronavirus. *EMBO J.*

- Isken, O., Grassmann, C. W., Sarisky, R. T., Kann, M., Zhang, S., Grosse, F., Kao, P. N. and Behrens, S.-E.** 2003. Members of the NF90/NFAR protein group are involved in the life cycle of a positive-strand RNA virus. *EMBO J.* **22**, 5655-5665.
- Izeta, A.** 2000. Replicación y encapsidación de minigenomas RNA derivados de coronavirus. expresión de genes heterólogos. Tesis doctoral. *Facultad de Ciencias. Departamento de Biología Molecular.* Universidad Autónoma de Madrid.
- Izeta, A., Smerdou, C., Alonso, S., Penzes, Z., Méndez, A., Plana-Durán, J. and Enjuanes, L.** 1999. Replication and packaging of transmissible gastroenteritis coronavirus-derived synthetic minigenomes. *J. Virol.* **73**, 1535-1545.
- Jacobs, L., Spaan, W. J. M., Horzinek, M. C. and van der Zeijst, B. A. M.** 1981. Synthesis of subgenomic mRNAs of mouse hepatitis virus is initiated independently: evidence from UV transcription mapping. *J. Virol.* **39**, 401.
- Jacobs, L., Van der Zeijst, B. A. M. and Horzinek, M.** 1986. Characterization and translation of transmissible gastroenteritis virus mRNAs. *J. Virol.* **57**, 1010-1015.
- Jacobson, A. and Peltz, S. W.** 1996. Interrelationships of the pathways of mRNA decay and translation in eukaryotic cells. *Annu. Rev. Biochem.* **65**, 693-739.
- Jaspars, E. M.** 1999. Genome activation in alfamo- and ilarviruses. *Arch. Virol.* **144**, 843-63.
- Jiménez, G., Correa, I., Melgosa, M. P., Bullido, M. J. and Enjuanes, L.** 1986. Critical epitopes in transmissible gastroenteritis virus neutralization. *J. Virol.* **60**, 131-139.
- Joo, M. and Makino, S.** 1995. Analysis of coronavirus transcription regulation. *Adv. Exp. Med. Biol.* **380**, 473-8.
- Kamitani, W., Narayanan, K., Huang, C., Lokugamage, K., Ikegami, T., Ito, N., Kubo, H. and Makino, S.** 2006. Severe acute respiratory syndrome coronavirus nsp1 protein suppresses host gene expression by promoting host mRNA degradation. *Proc. Natl. Acad. Sci. USA* **103**, 12885-90.
- Kampmueller, K. M. and Miller, D. J.** 2005. The cellular chaperone heat shock protein 90 facilitates Flock House virus RNA replication in *Drosophila* cells. *J. Virol.* **79**, 6827-37.
- Khromykh, A. A. and Westaway, E. G.** 1994. Completion of Kunjin virus RNA sequence and recovery of an infectious RNA transcribed from stably cloned full-length cDNA. *J. Virol.* **68**, 4580-4588.
- Khromykh, A. A. and Westaway, E. G.** 1997. Subgenomic replicons of the flavivirus kunjin: construction and applications. *J. Virol.* **71**, 1479-1505.
- Kim, Y. N., Lai, M. M. C. and Makino, S.** 1993. Generation and selection of coronavirus defective interfering RNA with large open reading frame by RNA recombination and possible editing. *Virology* **194**, 244-253.
- King, B. and Brian, D. A.** 1982. Bovine coronavirus structural proteins. *J. Virol.* **42**, 700-707.
- Krishnan, R., Chang, R. Y. and Brian, D. A.** 1996. Tandem placement of a coronavirus promoter results in enhanced mRNA synthesis from the downstream-most initiation site. *Virology* **218**, 400-405.

- Kuo, L., Godeke, G.-J., Raamsman, M. J. B., Masters, P. S. and Rottier, P. J. M.** 2000. Retargeting of coronavirus by substitution of the spike glycoprotein ectodomain: crossing the host cell species barrier. *J. Virol.* **74**, 1393-1406.
- Laemmli, U. K.** 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* **227**, 680-685.
- Lai, M. M. C.** 1986. Coronavirus leader-RNA-primed transcription: an alternative mechanism to RNA splicing. *BioEssays* **5**, 257-260.
- Lai, M. M. C.** 1998. Cellular factors in the transcription and replication of viral RNA genomes: a parallel to DNA-dependent RNA transcription. *Virology* **244**, 1-12.
- Lai, M. M. C. and Cavanagh, D.** 1997. The molecular biology of coronaviruses. *Adv. Virus Res.* **48**, 1-100.
- Lai, M. M. C. and Holmes, K. V.** 2001. Coronaviridae: The viruses and their replication. In *Fields Virology*, vol. 1 eds. D. M. Knipe and P. M. Howley), pp. 1163-1185. Philadelphia: Lippincott Williams & Wilkins.
- Law, A. H., Lee, D. C., Cheung, B. K., Yim, H. C. and Lau, A. S.** 2007. A Role for the Non-structural Protein-1 of SARS-Coronavirus in Chemokine Dysregulation. *J. Virol.* **81**, 416-22.
- Leparc-Goffart, I., Hingley, S. T., Chua, M. M., Jiang, X., Lavi, E. and Weiss, S. R.** 1997. Altered pathogenesis of a mutant of the murine coronavirus MHV-A59 is associated with a Q159L amino acid substitution in the spike protein. *Virology* **239**, 1-10.
- Li, H. P., Huang, P., Park, S. and Lai, M. M. C.** 1999. Polypyrimidine tract-binding protein binds to the leader RNA of mouse hepatitis virus and serves as a regulator of viral transcription. *J. Virol.* **73**, 772-777.
- Li, H. P., Zhang, X., Duncan, R., Comai, L. and Lai, M. M. C.** 1997. Heterogeneous nuclear ribonucleoprotein A1 binds to the transcription-regulatory region of mouse hepatitis virus RNA. *Proc. Natl. Acad. Sci. USA* **94**, 9544-9549.
- Liao, C.-L. and Lai, M. M. C.** 1995. A *cis*-acting viral protein is not required for the replication of a coronavirus defective-interfering RNA. *Virology* **209**, 428-436.
- Liljestrom, P.** 1994. Alphavirus expression systems. *Curr. Opin. Biotech.* **5**, 495-500.
- Lim, K. P. and Liu, D. X.** 1998. Characterisation of a papain-like proteinase domain encoded by ORF 1a of the coronavirus IBV and determination of the C-terminal cleavage site of an 87 kDa protein. *Adv. Exp. Med. Biol.* **440**, 173-184.
- Lin, Y. J. and Lai, M. M. C.** 1993. Deletion mapping of a mouse hepatitis virus defective interfering RNA reveals the requirement of an internal and discontinuous sequence for replication. *J. Virol.* **67**, 6110-6118.
- Lin, Y. J., Liao, C. L. and Lai, M. M. C.** 1994. Identification of the *cis*-acting signal for minus-strand RNA synthesis of a murine coronavirus: implications for the role of minus-strand RNA in RNA replication and transcription. *J. Virol.* **68**, 8131-8140.
- Lohmann, V., Korner, F., Koch, J., Herian, U., Theilmann, L. and Bartenschlager, R.** 1999. Replication of subgenomic hepatitis C virus RNAs in a hepatoma cell line. *Science* **285**, 110-3.

- Lomniczi, B. and Morser, J.** 1981. Polypeptides of infectious bronchitis virus. I. Polypeptides of the virion. *J. Gen. Virol.* **55**, 155-64.
- Luo, H., Chen, J., Chen, K., Shen, X. and Jiang, H.** 2006. Carboxyl terminus of severe acute respiratory syndrome coronavirus nucleocapsid protein: self-association analysis and nucleic acid binding characterization. *Biochemistry* **45**, 11827-35.
- Luytjes, W., Gerritsma, H. and Spaan, W. J. M.** 1996. Replication of synthetic defective interfering RNAs derived from coronavirus mouse hepatitis virus-A59. *Virology* **216**, 174-183.
- Makino, S., Yokomori, K. and Lai, M. M. C.** 1990. Analysis of efficiently packaged defective interfering RNAs of murine coronavirus: localization of a possible RNA-packaging signal. *J. Virol.* **64**, 6045-6053.
- Masters, P. S.** 1992. Localization of an RNA-binding domain in the nucleocapsid protein of the coronavirus mouse hepatitis virus. *Arch. Virol.* **125**, 141-160.
- Masters, P. S.** 1999. Reverse genetics of the largest RNA viruses. *Adv. Virus Res.* **53**, 245-264.
- Mathews, D. H., Sabina, J., Zuker, M. and Turner, D. H.** 1999. Expanded sequence dependence of thermodynamic parameters improves prediction of RNA secondary structure. *J. Mol. Biol.* **288**, 911-940.
- Mayer, M. P.** 2005. Recruitment of Hsp70 chaperones: a crucial part of viral survival strategies. *Rev. Physiol. Biochem. Pharmacol.* **153**, 1-46.
- Michel, Y. M., Borman, A. M., Paulous, S. and Kean, K. M.** 2001. Eukaryotic initiation factor 4G-poly(A) binding protein interaction is required for poly(A) tail-mediated stimulation of picornavirus internal ribosome entry segment-driven translation but not for X-mediated stimulation of hepatitis C virus translation. *Mol. Cell. Biol.* **21**, 4097-109.
- Minskaia, E., Hertzog, T., Gorbalenya, A. E., Campanacci, V., Cambillau, C., Canard, B. and Ziebuhr, J.** 2006. Discovery of an RNA virus 3'->5' exoribonuclease that is critically involved in coronavirus RNA synthesis. *Proc. Natl. Acad. Sci. USA* **103**, 5108-5113.
- Mislick, K. A. and Baldeschwieler, J. D.** 1996. Evidence for the role of proteoglycans in cation-mediated gene transfer. *Proc. Natl. Acad. Sci. USA* **93**, 12349-54.
- Molenkamp, R. and Spaan, W. J. M.** 1997. Identification of a specific interaction between the coronavirus mouse hepatitis virus A59 nucleocapsid protein and packaging signal. *Virology* **239**, 78-86.
- Molenkamp, R., van Tol, H., Rozier, B. C., van der Meer, Y., Spaan, W. J. and Snijder, E. J.** 2000. The arterivirus replicase is the only viral protein required for genome replication and subgenomic mRNA transcription. *J. Gen. Virol.* **81**, 2491-2496.
- Nanda, S. K., Johnson, R. F., Liu, Q. and Leibowitz, J. L.** 2004. Mitochondrial HSP70, HSP40, and HSP60 bind to the 3' untranslated region of the Murine hepatitis virus genome. *Arch. Virol.* **149**, 93-111.
- Nanda, S. K. and Leibowitz, J. L.** 2001. Mitochondrial aconitase binds to the 3' untranslated region of the mouse hepatitis virus genome. *J. Virol.* **75**, 3352-3362.

- Narayanan, K. and Makino, S.** 2001. Cooperation of an RNA packaging signal and a viral envelope protein in coronavirus RNA Packaging. *J. Virol.* **75**, 9059-9067.
- Nelson, G. W. and Stohlman, S. A.** 1993. Localization of the RNA-binding domain of mouse hepatitis virus nucleocapsid protein. *J. Gen. Virol.* **74**, 1975-1979.
- Nelson, G. W., Stohlman, S. A. and Tahara, S. M.** 2000. High affinity interaction between nucleocapsid protein and leader/intergenic sequence of mouse hepatitis virus RNA. *J. Gen. Virol.* **81**, 181-188.
- Norman, J. O., McClurkin, A. W. and Bachrach, H. L.** 1968. Infectious nucleic acid from a transmissible agent causing gastroenteritis in pigs. *J. Comp. Pathol.* **78**, 227-35.
- Ortego, J., Escors, D., Laude, H. and Enjuanes, L.** 2002. Generation of a replication-competent, propagation-deficient virus vector based on the transmissible gastroenteritis coronavirus genome. *J. Virol.* **76**, 11518-11529.
- Ortego, J., Sola, I., Almazán, F., Ceriani, J. E., Riquelme, C., Balasch, M., Planadurán, J. and Enjuanes, L.** 2003. Transmissible gastroenteritis coronavirus gene 7 is not essential but influences *in vivo* virus replication and virulence. *Virology* **308**, 13-22.
- Otero, L. J., Ashe, M. P. and Sachs, A. B.** 1999. The yeast poly(A)-binding protein Pab1p stimulates *in vitro* poly(A)-dependent and cap-dependent translation by distinct mechanisms. *EMBO J.* **18**, 3153-3163.
- Ozdarendeli, A., Ku, S., Rochat, S., Williams, G. D., Senanayake, S. D. and Brian, D. A.** 2001. Downstream sequences influence the choice between a naturally occurring noncanonical and closely positioned upstream canonical heptameric fusion motif during bovine coronavirus subgenomic mRNA synthesis. *J. Virol.* **75**, 7362-7374.
- Page, K. W., Britton, P. and Bournsnel, M. E. G.** 1990. Sequence analysis of the leader RNA of two porcine coronaviruses: transmissible gastroenteritis coronavirus and porcine respiratory coronavirus. *Virus Genes* **4**, 289-301.
- Parsley, T. B., Towner, J. S., Blyn, L. B., Ehrenfeld, E. and Semler, B. L.** 1997. Poly (rC) binding protein 2 forms a ternary complex with the 5'-terminal sequences of poliovirus RNA and the viral 3CD proteinase. *RNA* **3**, 1124-34.
- Pensaert, M. and Van Reeth, K.** 1998. Porcine epidemic diarrhea and porcine respiratory coronavirus. *Proc. Ame. Ass. Swi. Pract.*, 433-436.
- Perlman, S., Lane, T. E. and Buchmeier, M. J.** 2000. Coronavirus: hepatitis, peritonitis, and central nervous system disease. In *Effects of microbes on the immune system*, eds. M. W. Cunningham and R. S. Fujinami, pp. 331-348. Philadelphia: Lippincott Williams and Wilkins.
- Piron, M., Vende, P., Cohen, J. and Poncet, D.** 1998. Rotavirus RNA-binding protein NSP3 interacts with eIF4GI and evicts the poly(A) binding protein from eIF4F. *EMBO J.* **17**, 5811-5821.
- Pushko, P., Parker, M., Ludwing, G. V., Davis, N. L., Johnston, R. E. and Smith, J. F.** 1997. Replication-helper systems from attenuated Venezuelan equine encephalitis virus: expression of heterologous genes *in vitro* and immunization against heterologous pathogens *in vivo*. *Virology* **239**, 389-401.

- Putics, A., Slaby, J., Filipowicz, W., Gorbalenya, A. E. and Ziebuhr, J.** 2006. ADP-ribose-1"-phosphatase activities of the human coronavirus 229E and SARS coronavirus X domains. *Adv. Exp. Med. Biol.* **581**, 93-6.
- Raman, S., Bouma, P., Williams, G. D. and Brian, D. A.** 2003. Stem-loop III in the 5' untranslated region is a cis-acting element in bovine coronavirus defective interfering RNA replication. *J. Virol.* **77**, 6720-30.
- Raman, S. and Brian, D. A.** 2005. Stem-loop IV in the 5' untranslated region is a cis-acting element in bovine coronavirus defective interfering RNA replication. *J. Virol.* **79**, 12434-46.
- Renzi, F., Caffarelli, E., Laneve, P., Bozzoni, I., Brunori, M. and Vallone, B.** 2006. The structure of the endoribonuclease XendoU: From small nucleolar RNA processing to severe acute respiratory syndrome coronavirus replication. *Proc. Natl. Acad. Sci. USA* **103**, 12365-70.
- Repass, J. F. and Makino, S.** 1998. Importance of the positive-strand RNA secondary structure of a murine coronavirus defective interfering RNA internal replication signal in positive-strand RNA synthesis. *J. Virol.* **72**, 7926-7933.
- Ricagno, S., Egloff, M. P., Ulferts, R., Coutard, B., Nurizzo, D., Campanacci, V., Cambillau, C., Ziebuhr, J. and Canard, B.** 2006. Crystal structure and mechanistic determinants of SARS coronavirus nonstructural protein 15 define an endoribonuclease family. *Proc. Natl. Acad. Sci. USA* **103**, 11892-7.
- Rota, P. A., Oberste, M. S., Monroe, S. S., Nix, W. A., Campganoli, R., Icenogle, J. P., Peñaranda, S., Bankamp, B., Maher, K., Chen, M.-H. et al.** 2003. Characterization of a novel coronavirus associated with severe acute respiratory syndrome. *Science* **300**, 1394-1399.
- Rowland, R. R., Chauhan, V., Fang, Y., Pekosz, A., Kerrigan, M. and Burton, M. D.** 2005. Intracellular localization of the severe acute respiratory syndrome coronavirus nucleocapsid protein: absence of nucleolar accumulation during infection and after expression as a recombinant protein in vero cells. *J. Virol.* **79**, 11507-11512.
- Ryzhova, E. V., Vos, R. M., Albright, A. V., Harrist, A. V., Harvey, T. and Gonzalez-Scarano, F.** 2006. Annexin 2: a novel human immunodeficiency virus type 1 Gag binding protein involved in replication in monocyte-derived macrophages. *J. Virol.* **80**, 2694-704.
- Sachs, A. B. and Varani, G.** 2000. Eukaryotic translation initiation: there are (at least) two sides to every story. *Nat. Struct. Biol.* **7**, 356-61.
- Saif, L. J. and Wesley, R. D.** 1992. Transmissible gastroenteritis. In *Diseases of Swine*, eds. A. D. Leman B. E. Straw W. L. Mengeling S. D'Allaire and D. J. Taylor, pp. 362-386. Ames. Iowa: Wolfe Publishing Ltd.
- Salanueva, I. J., Carrascosa, J. L. and Risco, C.** 1999. Structural maturation of the transmissible gastroenteritis coronavirus. *J. Virol.* **73**, 7952-7964.
- Sambrook, J. and Russell, D. W.** 2001. Molecular cloning: A laboratory manual. Cold Spring Harbor, New York: Cold Spring Harbor Laboratory Press.

- Sánchez, C. M., Gebauer, F., Suñé, C., Mendez, A., Dopazo, J. and Enjuanes, L.** 1992. Genetic evolution and tropism of transmissible gastroenteritis coronaviruses. *Virology* **190**, 92-105.
- Sánchez, C. M., Izeta, A., Sánchez-Morgado, J. M., Alonso, S., Sola, I., Balasch, M., Plana-Duran, J. and Enjuanes, L.** 1999. Targeted recombination demonstrates that the spike gene of transmissible gastroenteritis coronavirus is a determinant of its enteric tropism and virulence. *J. Virol.* **73**, 7607-7618.
- Sánchez, C. M., Jiménez, G., Laviada, M. D., Correa, I., Suñé, C., Bullido, M. J., Gebauer, F., Smerdou, C., Callebaut, P., Escribano, J. M. et al.** 1990. Antigenic homology among coronaviruses related to transmissible gastroenteritis virus. *Virology* **174**, 410-417.
- Sawicki, S. G. and Sawicki, D. L.** 1990. Coronavirus transcription: subgenomic mouse hepatitis virus replicative intermediates function in RNA synthesis. *J. Virol.* **64**, 1050-1056.
- Sawicki, S. G. and Sawicki, D. L.** 2005. Coronavirus transcription: a perspective. *Curr. Top. Microbiol. Immunol.* **287**, 31-55.
- Sawicki, S. G., Sawicki, D. L. and Siddell, S. G.** 2007. A contemporary view of coronavirus transcription. *J. Virol.* **81**, 20-9.
- Schechter, I. and Berger, A.** 1967. On the size of the active site in proteases. I. Papain. *Biochem. Biophys. Res. Commun.* **27**, 157-62.
- Schelle, B., Karl, N., Ludewig, B., Siddell, S. G. and Thiel, V.** 2005. Selective replication of coronavirus genomes that express nucleocapsid protein. *J. Virol.* **79**, 6620-6630.
- Serva, S. and Nagy, P. D.** 2006. Proteomics analysis of the tombusvirus replicase: Hsp70 molecular chaperone is associated with the replicase and enhances viral RNA replication. *J. Virol.* **80**, 2162-9.
- Sethna, P. B., Hung, S.-L. and Brian, D. A.** 1989. Coronavirus subgenomic minus-strand RNAs and the potential for mRNA replicons. *Proc. Natl. Acad. Sci. USA* **86**, 5626-5630.
- Shen, X. and Masters, P. S.** 2001. Evaluation of the role of heterogeneous nuclear ribonucleoprotein A1 as a host factor in murine coronavirus discontinuous transcription and genome replication. *Proc. Natl. Acad. Sci. USA* **98**, 2717-2722.
- Shi, S. T., Huang, P., Li, H. P. and Lai, M. M. C.** 2000. Heterogeneous nuclear ribonucleoprotein A1 regulates RNA synthesis of a cytoplasmic virus. *EMBO J.* **19**, 4701-4711.
- Shi, S. T., Yu, G. Y. and Lai, M. M.** 2003. Multiple type A/B heterogeneous nuclear ribonucleoproteins (hnRNPs) can replace hnRNP A1 in mouse hepatitis virus RNA synthesis. *J. Virol.* **77**, 10584-10593.
- Siddell, S. G., Sawicki, D., Meyer, Y., Thiel, V. and Sawicki, S.** 2001. Identification of the mutations responsible for the phenotype of three MHV RNA-negative ts mutants. *Adv. Exp. Med. Biol.* **494**, 453-458.
- Silvestri, L. S., Parilla, J. M., Morasco, B. J., Ogram, S. A. and Flanagan, J. B.** 2006. Relationship between poliovirus negative-strand RNA synthesis and the length of the 3' poly(A) tail. *Virology* **345**, 509-19.

- Sims, A. C., Ostermann, J. and Denison, M. R.** 2000. Mouse hepatitis virus replicase proteins associate with two distinct populations of intracellular membranes. *J. Virol.* **74**, 5647-54.
- Singh, R. and Valcarcel, J.** 2005. Building specificity with nonspecific RNA-binding proteins. *Nat. Struct. Mol. Biol.* **12**, 645-53.
- Snijder, E. J., Bredenbeek, P. J., Dobbe, J. C., Thiel, V., Ziebuhr, J., Poon, L. L. M., Guan, Y., Rozanov, M., Spaan, W. J. M. and Gorbalenya, A. E.** 2003. Unique and conserved features of genome and proteome of SARS-coronavirus, an early split-off from the coronavirus group 2 lineage. *J. Mol. Biol.* **331**, 991-1004.
- Snijder, E. J. and Spaan, W. J. M.** 1995. The coronaviruslike superfamily. In *The Coronaviridae*, ed. S. G. Siddell, pp. 239-252. New York: Plenum press.
- Snijder, E. J., van der Meer, Y., Zevenhoven-Dobbe, J., Onderwater, J. J., van der Meulen, J., Koerten, H. K. and Mommaas, A. M.** 2006. Ultrastructure and origin of membrane vesicles associated with the severe acute respiratory syndrome coronavirus replication complex. *J. Virol.* **80**, 5927-5940.
- Sola, I., Alonso, S., Zúñiga, S., Balach, M., Plana-Durán, J. and Enjuanes, L.** 2003. Engineering transmissible gastroenteritis virus genome as an expression vector inducing latogenic immunity. *J. Virol.* **77**, 4357-4369.
- Sola, I., Moreno, J. L., Zúñiga, S., Alonso, S. and Enjuanes, L.** 2005. Role of nucleotides immediately flanking the transcription-regulating sequence core in coronavirus subgenomic mRNA synthesis. *J. Virol.* **79**, 2506-2516.
- Somogyi, P., Jenner, A. J., Brierley, I. and Inglis, S. C.** 1993. Ribosomal pausing during translation of an RNA pseudoknot. *Mol. Cell. Biol.* **13**, 6931-40.
- Spagnolo, J. F. and Hogue, B. G.** 2000. Host protein interactions with the 3' end of bovine coronavirus RNA and the requirement of the poly(A) tail for coronavirus defective genome replication. *J. Virol.* **74**, 5053-5065.
- St-Jean, J. R., Desforges, M., Almazán, F., Jacomy, H., Enjuanes, L. and Talbot, P. J.** 2006. Recovery of a neurovirulent human coronavirus OC43 from an infectious cDNA clone. *J. Virol.* **80**, 3670-3674.
- Stalcup, R. P., Baric, R. S. and Leibowitz, J. L.** 1998. Genetic complementation among three panels of mouse hepatitis virus gene 1 mutants. *Virology* **241**, 112-121.
- Stern, D. F. and Sefton, B. M.** 1982. Synthesis of coronavirus mRNAs: kinetics of inactivation of IBV RNA synthesis by UV light. *J. Virol.* **42**, 755.
- Stohlman, S. A., Baric, R. S., Nelson, G. N., Soe, L. H., Welter, L. M. and Deans, R. J.** 1988. Specific interaction between coronavirus leader RNA and nucleocapsid protein. *J. Virol.* **62**, 4288-4295.
- Stohlman, S. A. and Lai, M. M.** 1979. Phosphoproteins of murine hepatitis viruses. *J. Virol.* **32**, 672-675.
- Sturman, L. S., Holmes, K. V. and Behnke, J.** 1980. Isolation of coronavirus envelope glycoproteins and interaction with the viral nucleocapsid. *J. Virol.* **33**, 449-462.
- Sturman, L. S. and Holmes, K. V.** 1977. Characterization of a coronavirus. II. Glycoproteins of the viral envelope: tryptic peptide analysis. *Virology* **77**, 650-660.

- Sulea, T., Lindner, H. A., Purisima, E. O. and Menard, R.** 2006. Binding site-based classification of coronaviral papain-like proteases. *Proteins* **62**, 760-75.
- Svitkin, Y. V., Imataka, H., Khaleghpour, K., Kahvejian, A., Liebig, H. D. and Sonenberg, N.** 2001. Poly(A)-binding protein interaction with eIF4G stimulates picornavirus IRES-dependent translation. *RNA* **7**, 1743-52.
- Tahara, S. M., Dietlin, T. A., Bergmann, C. C., Nelson, G. W., Kyuwa, S., Anthony, R. P. and Stohlman, S. A.** 1994. Coronavirus translational regulation: leader affects mRNA efficiency. *Virology* **202**, 621-630.
- Tahara, S. M., Dietlin, T. A., Nelson, G. W., Stohlman, S. A. and Manno, D. J.** 1998. Mouse hepatitis virus nucleocapsid protein as a translational effector of viral mRNAs. *Adv. Exp. Med. Biol.* **440**, 313-318.
- Tang, M. X., Redemann, C. T. and Szoka, F. C., Jr.** 1996. In vitro gene delivery by degraded polyamidoamine dendrimers. *Bioconj. Chem.* **7**, 703-14.
- Tarun, S. Z., Jr. and Sachs, A. B.** 1996. Association of the yeast poly(A) tail binding protein with translation initiation factor eIF-4G. *EMBO J.* **15**, 7168-77.
- Teng, H., Piñon, J. D. and Weiss, S. R.** 1999. Expression of murine coronavirus recombinant papain-like proteinase: efficient cleavage is dependent on the lengths of both the substrate and the proteinase polypeptides. *J. Virol.* **73**, 2658-2666.
- Terribilini, M., Lee, J. H., Yan, C., Jernigan, R. L., Honavar, V. and Dobbs, D.** 2006. Prediction of RNA binding sites in proteins from amino acid sequence. *RNA* **12**, 1450-62.
- Thiel, V., Herold, J., Schelle, B. and Siddell, S.** 2001a. Infectious RNA transcribed *in vitro* from a cDNA copy of the human coronavirus genome cloned in vaccinia virus. *J. Gen. Virol.* **82**, 1273-1281.
- Thiel, V., Herold, J., Schelle, B. and Siddell, S. G.** 2001b. Viral replicase gene products suffice for coronavirus discontinuous transcription. *J. Virol.* **75**, 6676-6681.
- Tijms, M. A., van Dinten, L. C., Gorbalenya, A. E. and Snijder, E. J.** 2001. A zinc finger-containing papain-like protease couples subgenomic mRNA synthesis to genome translation in a positive-stranded RNA virus. *Proc. Natl. Acad. Sci. USA* **98**, 1889-1894.
- van der Hoek, L., Pyrc, K., Jebbink, M. F., Vermeulen-Oost, W., Berkhout, R. J., Wolthers, K. C., Wertheim-van Dillen, P. M., Kaandorp, J., Spaargaren, J. and Berkhout, B.** 2004. Identification of a new human coronavirus. *Nat. Med.* **10**, 368-373.
- van der Meer, Y., Snijder, E. J., Dobbe, J. C., Schleich, S., Denison, M. R., Spaan, W. J. M. and Krijnse Locker, J.** 1999. Localization of mouse hepatitis virus nonstructural proteins and RNA synthesis indicates a role for late endosomes in viral replication. *J. Virol.* **73**, 7641-7657.
- van der Most, R., Luytjes, W., Rutjes, S. and Spaan, W. J. M.** 1995. Translation but not the encoded sequence is essential for the efficient propagation of the defective interfering RNAs of the coronavirus mouse hepatitis virus. *J. Virol.* **69**, 3744-3751.

- van der Most, R. G., Bredenbeek, P. J. and Spaan, W. J. M.** 1991. A domain at the 3' end of the polymerase gene is essential for encapsidation of coronavirus defective interfering RNAs. *J. Virol.* **65**, 3219-3226.
- van der Vossen, E. A., Neeleman, L. and Bol, J. F.** 1994. Early and late functions of alfalfa mosaic virus coat protein can be mutated separately. *Virology* **202**, 891-903.
- van Marle, G., Dobbe, J. C., Gultyaev, A. P., Luytjes, W., Spaan, W. J. M. and Snijder, E. J.** 1999. Arterivirus discontinuous mRNA transcription is guided by base pairing between sense and antisense transcription-regulating sequences. *Proc. Natl. Acad. Sci. USA* **96**, 12056-12061.
- von Grotthuss, M., Wyrwicz, L. S. and Rychlewski, L.** 2003. mRNA cap-1 methyltransferase in the SARS genome. *Cell* **113**, 701-702.
- Waggoner, S. and Sarnow, P.** 1998. Viral ribonucleoprotein complex formation and nucleolar-cytoplasmic relocation of nucleolin in poliovirus-infected cells. *J. Virol.* **72**, 6699-709.
- Walker, P. J., Bonami, J. R., Boonsaeng, V., Chang, P. S., Cowley, J. A., Enjuanes, L., Flegel, T. W., Lightner, D. V., Loh, P. C., Snijder, E. J. et al.** 2005. *Roniviridae*. In *Virus Taxonomy. Eighth Report of the International Committee on Taxonomy of Viruses*, eds. C. M. Fauquet M. A. Mayo J. Maniloff U. Desselberger and L. A. Ball, pp. 975-979. San Diego, California: Academic Press.
- Walter, B. L., Parsley, T. B., Ehrenfeld, E. and Semler, B. L.** 2002. Distinct poly(rC) binding protein KH domain determinants for poliovirus translation initiation and viral RNA replication. *J. Virol.* **76**, 12008-22.
- Wang, K., Boysen, C., Shizuya, H., Simon, M. I. and Hood, L.** 1997. Complete nucleotide sequence of two generations of a bacterial artificial chromosome cloning vector. *BioTechniques* **23**, 992-994.
- Wang, Y. and Zhang, X.** 1999. The nucleocapsid protein of coronavirus mouse hepatitis virus interacts with the cellular heterogeneous nuclear ribonucleoprotein A1 in vitro and in vivo. *Virology* **265**, 96-109.
- Waterston, R. H. Lindblad-Toh, K. Birney, E. Rogers, J. Abril, J. F. Agarwal, P. Agarwala, R. Ainscough, R. Alexandersson, M. An, P. et al.** 2002. Initial sequencing and comparative analysis of the mouse genome. *Nature* **420**, 520-62.
- Wernersson, R., Schierup, M. H., Jorgensen, F. G., Gorodkin, J., Panitz, F., Staerfeldt, H. H., Christensen, O. F., Mailund, T., Hornshøj, H., Klein, A. et al.** 2005. Pigs in sequence space: a 0.66X coverage pig genome survey based on shotgun sequencing. *BMC Genomics* **6**, 70.
- Williams, G. D., Chang, R.-Y. and Brian, D. A.** 1999. A phylogenetically conserved hairpin-type 3' untranslated region pseudoknot functions in coronavirus RNA replication. *J. Virol.* **73**, 8349-8355.
- Wilson, L., Gage, P. and Ewart, G.** 2006. Validation of coronavirus E proteins ion channels as targets for antiviral drugs. In *The Nidovirus: Towards control of SARS and other nidovirus diseases*, vol. 581 eds. S. Perlman and K. Holmes, pp. 573-578. New York: Springer.
- Wilson, L., McKinlay, C., Gage, P. and Ewart, G.** 2004. SARS coronavirus E protein forms cation-selective ion channels. *Virology* **330**, 322-331.

- Woo, P. C., Lau, S. K., Chu, C. M., Chan, K. H., Tsoi, H. W., Huang, Y., Wong, B. H., Poon, R. W., Cai, J. J., Luk, W. K. et al.** 2005. Characterization and complete genome sequence of a novel coronavirus, coronavirus HKU1, from patients with pneumonia. *J. Virol.* **79**, 884-895.
- Woods, R. D. and Wesley, R. D.** 1998. Transmissible gastroenteritis coronavirus carrier sow. *Adv. Exp. Med. Biol.* **440**, 641-648.
- Wurm, T., Chen, H., Hodgson, T., Britton, P., Brooks, G. and Hiscox, J. A.** 2001. Localization to the nucleolus is a common feature of coronavirus nucleoproteins, and the protein may disrupt host cell division. *J. Virol.* **75**, 9345-9356.
- You, S., Falgout, B., Markoff, L. and Padmanabhan, R.** 2001. In vitro RNA synthesis from exogenous dengue viral RNA templates requires long range interactions between 5'- and 3'-terminal regions that influence RNA structure. *J. Biol. Chem.* **276**, 15581-91.
- Yount, B., Curtis, K. M. and Baric, R. S.** 2000. Strategy for systematic assembly of large RNA and DNA genomes: the transmissible gastroenteritis virus model. *J. Virol.* **74**, 10600-10611.
- Yount, B., Curtis, K. M., Fritz, E. A., Hensley, L. E., Jahrling, P. B., Prentice, E., Denison, M. R., Geisbert, T. W. and Baric, R. S.** 2003. Reverse genetics with a full-length infectious cDNA of severe acute respiratory syndrome coronavirus. *Proc. Natl. Acad. Sci. USA* **100**, 12995-13000.
- Yount, B., Denison, M. R., Weiss, S. R. and Baric, R. S.** 2002. Systematic assembly of a full length infectious cDNA of mouse hepatitis virus strain A59. In *21st Annual Meeting*, (ed. ASV), pp. 109. Lexington, Kentucky: American Society for Virology.
- Yu, I. M., Gustafson, C. L., Diao, J., Burgner, J. W., 2nd, Li, Z., Zhang, J. and Chen, J.** 2005. Recombinant severe acute respiratory syndrome (SARS) coronavirus nucleocapsid protein forms a dimer through its C-terminal domain. *J. Biol. Chem.* **280**, 23280-6.
- Zakhartchouk, A. N., Viswanathan, S., Mahony, J. B., Gauldie, J. and Babiuk, L. A.** 2005. Severe acute respiratory syndrome coronavirus nucleocapsid protein expressed by an adenovirus vector is phosphorylated and immunogenic in mice. *J. Gen. Virol.* **86**, 211-215.
- Zhang, X. and Lai, M. M. C.** 1995. Interactions between the cytoplasmic proteins and the intergenic (promoter) sequence of mouse hepatitis virus RNA: correlation with the amounts of subgenomic mRNA transcribed. *J. Virol.* **69**, 1637-1644.
- Zhou, M., Williams, A. K., Chung, S. I., Wang, L. and Collisson, E. W.** 1996. The infectious bronchitis virus nucleocapsid protein binds RNA sequences in the 3' terminus of the genome. *Virology* **217**, 191-199.
- Ziebuhr, J.** 2005. The coronavirus replicase. In *Coronavirus replication and reverse genetics*, vol. 287 (ed. L. Enjuanes), pp. 57-94: Springer.
- Ziebuhr, J., Snijder, E. J. and Gorbalenya, A. E.** 2000. Virus-encoded proteinases and proteolytic processing in the *Nidovirales*. *J. Gen. Virol.* **81**, 853-879.

- Ziebuhr, J., Thiel, V. and Gorbalenya, A. E.** 2001. The autocatalytic release of a putative RNA virus transcription factor from its polyprotein precursor involves two paralogous papain-like proteases that cleave the same peptide bond. *J. Biol. Chem.* **276**, 33220-33232.
- Zuker, M.** 2003. Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic Acids Res.* **31**, 3406-3415.
- Zúñiga, S., Sola, I., Alonso, S. and Enjuanes, L.** 2004. Sequence motifs involved in the regulation of discontinuous coronavirus subgenomic RNA synthesis. *J. Virol.* **78**, 980-994.
- Zúñiga, S., Sola, I., Moreno, J. L., Sabella, P., Plana-Durán, J. and Enjuanes, L.** 2007. Coronavirus nucleocapsid protein is an RNA chaperone. *Virology* **357**, 215-27.

X. ANEXO

The Nucleoprotein Is Required for Efficient Coronavirus Genome Replication

Fernando Almazán, Carmen Galán, and Luis Enjuanes*

Department of Molecular and Cell Biology, Centro Nacional de Biotecnología, Consejo Superior de Investigaciones Científicas, Campus Universidad Autónoma, Cantoblanco, Madrid, Spain

Received 13 March 2004/Accepted 14 June 2004

The construction of a set of transmissible gastroenteritis coronavirus (TGEV)-derived replicons as bacterial artificial chromosomes is reported. These replicons were generated by sequential deletion of nonessential genes for virus replication, using a modified TGEV full-length cDNA clone containing unique restriction sites between each pair of consecutive genes. Efficient activity of TGEV replicons was associated with the presence of the nucleoprotein provided either in *cis* or in *trans*. TGEV replicons were functional in several cell lines, including the human cell line 293T, in which no or very low cytopathic effect was observed, and expressed high amounts of heterologous protein.

Since the etiologic pathogen causing severe acute respiratory syndrome was identified as a coronavirus (6, 8, 12, 15, 16, 18, 25), the study of coronavirus molecular biology has acquired significant relevance in order to develop effective strategies to prevent and control coronavirus infections. The design of antiviral drugs interfering with coronavirus replication is a rational approach requiring a detailed study of the replication mechanism at the molecular level. Coronavirus-derived replicons should be useful tools to select effective interfering molecules.

Transmissible gastroenteritis coronavirus (TGEV) is a member of the *Coronaviridae* family within the order *Nidovirales*, composed of enveloped single-stranded, positive-sense RNA viruses relevant in animal and human health (7). About two-thirds of the 28.5-kb TGEV genome encodes the replicase gene, which comprises open reading frames 1a and 1b, the last one being expressed by ribosomal frameshifting (23). Translation of both open reading frames results in the synthesis of two large polypeptides that are processed by viral proteinases to yield the replicase-transcriptase complex (38). The 3' one-third of the genome includes the genes encoding the structural and nonstructural proteins, in the order 5'-S-3a-3b-E-M-N-7-3'. These proteins are expressed by a discontinuous transcription process that most probably takes place during the synthesis of the negative strand, leading to the generation of a 3' coterminal nested set of subgenomic mRNAs, each of which has at its 5' end a capped leader sequence derived from the 5' end of the genome (26, 39). Synthesis of subgenomic negative-sense RNA species is regulated by the transcription-regulating sequences (TRSs), which include a highly conserved core sequence that is found preceding each gene and at the 3' end of the leader sequence (3).

Little is known about the genome replication of coronavirus at the molecular level, and until recently, study was restricted to the analysis of defective interfering RNA genomes that are

amplified by a helper virus (13, 24, 34), temperature-sensitive mutants that are defective in RNA synthesis (27, 29), and the analysis of recombinant viruses generated by targeted recombination (19). The recent construction of coronavirus full-length cDNA clones (2, 4, 30, 35–37) provides an opportunity for the genetic manipulation of coronavirus genomes to study fundamental viral processes and to develop expression vectors. In fact, the generation of coronavirus self-replicating RNAs (replicons) provides a valuable avenue to explore the molecular bases of coronavirus genome replication. This approach has been used to analyze the replication and transcription mechanism of many positive-strand RNA viruses (1, 11, 17, 20, 31, 33).

In this article, we report the generation and analysis of a collection of TGEV-derived replicons that are functional in several cell lines and express high amounts of heterologous proteins. In addition, our data indicate that the N protein is required either in *cis* or in *trans* for efficient coronavirus-based replicon activity.

Strategy for construction of TGEV-derived replicons. The generation of a TGEV replicon was based in a first approach on published data showing that for many positive-strand RNA viruses, only the replicase gene was needed for autonomous replication of the viral RNA (9, 14, 17, 20). Accordingly, a cDNA that contains the 5' and 3' ends of the TGEV genome and the replicase gene was cloned as a bacterial artificial chromosome (BAC) under the control of the cytomegalovirus (CMV) immediate-early promoter following the same approach as that described for the generation of a TGEV full-length cDNA clone (2, 10). Unfortunately, after transfection of swine testis (ST) cells with the replicon, using Lipofectine (Invitrogen), no significative replicase activity was detected (data not shown). One possible explanation for this lack of activity could be the low transfection efficiency of ST cells (less of 0.1%) with large-size coronavirus replicons. To test this possibility, the transfection efficiency of TGEV BAC clones was optimized. Baby hamster kidney cells stably transformed with the gene encoding the porcine aminopeptidase N (BHK-pAPN) (5) and human 293T cells (American Type Culture Collection) were used in the optimization process, as they are

* Corresponding author. Mailing address: Department of Molecular and Cell Biology, Centro Nacional de Biotecnología, CSIC, Campus Universidad Autónoma, Cantoblanco, 28049 Madrid, Spain. Phone: 34 91 585 4555. Fax: 34 91 585 4915. E-mail: L.Enjuanes@cnb.uam.es.

transfected with efficiencies about 10- and 20-fold higher, respectively, than ST cells and because they also are permissive for TGEV replication. After more than 12 commercial transfection reagents were tested, Lipofectamine 2000 (Invitrogen) was selected for transfection of BACs containing TGEV-derived replicon cDNA clones (about 30 kb), leading to transfection efficiencies of 20 and 42% in BHK-pAPN and 293T cells, respectively (F. Almazán, C. Galán, and L. Enjuanes, unpublished data).

After optimization of BAC transfection, a collection of TGEV replicons was generated to study the roles of other viral genes, in addition to the replicase gene, on coronavirus replicon activity. Genes 3a, 3b, and 7 are nonessential for virus replication in culture cells, and gene E, although essential for virus production, does not affect genome replication (21, 22, 28). A modified TGEV infectious cDNA clone (TGEV-RS), in which overlapping genes were separated by duplication of TRSs and introduction of unique restriction sites at the 5' end of each gene (22), was used for the construction of three TGEV replicons by sequential deletion of structural and nonstructural genes nonessential for virus genome replication: (i) REP 1, containing the untranslated 5' and 3' regions of the genome, the replicase, M, N, and 7 genes; (ii) REP 2, in which gene M was deleted; (iii) and REP 3, with genes M and N deleted (Fig. 1A). Gene 7 was maintained in all cases because it may contain necessary *cis* elements for genome replication due to its proximity to the 3' end of the TGEV genome. Expression of gene 7 mRNA was used to study replicon functionality by reverse transcription-PCR (RT-PCR) analysis. Additionally, a multicloning site containing unique restriction sites, *AvrII*, *MluI*, *SwaI*, and *FseI*, was cloned downstream of the replicase gene to allow cloning and expression of heterologous genes. The detailed cloning strategy, plasmid maps, and sequences are available from the authors upon request.

Activity of TGEV-derived replicons in several cell lines. To study the functionality of TGEV replicons, human 293T cells, which are transfected with efficiencies over 40%, were used in the assay. 293T cells were grown to 95% confluence on 35-mm-diameter plates and transfected with 4 μ g of either TGEV replicons, the full-length cDNA clone, or a nonreplicative cDNA clone, using 12 μ g of Lipofectamine 2000 (Invitrogen) according to the manufacturer's specifications. Total intracellular RNA was extracted at 1.5, 5, and 10 days posttransfection (dpt) with the RNeasy Mini Kit (QIAGEN) and used as the template for RT-PCR analysis of gene 7 mRNA transcription. RT reactions were performed with Moloney murine leukemia virus reverse transcriptase (Ambion) and the antisense primer 7(213)RS (5'-TCTGTAGCAGCAAAATCC-3'), complementary to nucleotides 199 to 216 of gene 7. The cDNAs generated were amplified by PCR using the reverse primer 7(213)RS and the forward primer SP (5'-GTGAGTGTAGCGTGGCTATATCTCTTC-3'), spanning nucleotides 15 to 41 of the TGEV leader sequence. High levels of gene 7 mRNA were detected in cells transfected with REP 1 and 2 (Fig. 1B), indicating that both replicons are functional, at least at the transcription level. However, no transcript was detected in cells transfected with REP 3, lacking the N gene, indicating a possible role of N protein in replicon activity (Fig. 1B). Interestingly, gene 7 transcripts were detected in REP 1 and 2 even after 10 dpt

(three cell passages), suggesting that these replicons had very little cytopathicity or were noncytopathic in human cells. Identical experiments were performed with BHK-pAPN and ST cells, but in these cases TGEV replicons were cytopathic and gene 7 mRNA was detected only in the samples analyzed up to 36 h posttransfection (hpt) but not later (data not shown). A quantitative analysis of gene 7 mRNA in transfected 293T, BHK-pAPN, and ST cells was performed by real-time RT-PCR. The primers used for RT (7RS, 5'-AAACTGTAATAAATACAGCATGGAGGAA-3', complementary to nucleotides 10 to 38 of gene 7) and PCR (reverse primer 7RS and the forward primer LDVS, 5'-CGTGGCTATATCTCTTCTTTTCTTTAACTAG-3', spanning nucleotides 24 to 55 of the TGEV leader sequence) were designed by using the Primer Express software (Applied Biosystems). SYBR Green PCR master mix was used in the PCR step following the manufacturer's specifications (Applied Biosystems). In 293T cells the amount of transcript was practically identical in the samples analyzed at 1.5 and 10 dpt, in contrast to those in BHK-pAPN and ST cells, in which transcript expression was undetectable after 2 dpt, reinforcing the idea that TGEV replicons are noncytopathic in human 293T cells. As the TGEV replication level was about threefold lower in 293T cells than in BHK-pAPN cells (data not shown), this could explain the fact that TGEV replicons were cytopathic in BHK-pAPN cells and not in 293T cells. However, replicon toxicity could also be due to a cell-specific factor. Finally, the amount of the cDNA encoding the replicon RNA and the RNA itself were analyzed at different dpt. Replicon RNA showed a higher stability than the cDNA (data not shown), suggesting that the TGEV replicon RNA was also functional at the replication level.

Role of N protein provided either in *cis* or in *trans* in replicon activity. To confirm whether the N protein is important for TGEV replicon activity, the synthesis of gene 7 mRNA was analyzed by RT-PCR in BHK-pAPN cells transformed with the Sindbis virus replicon pSINrep21 (9) expressing TGEV N protein. To show the specificity of N protein expression, similar analyses were carried out with BHK-pAPN cells and BHK-pAPN cells transformed with the Sindbis virus replicon alone or with the replicon expressing the TGEV E protein (Fig. 2A and B). The indicated cells were grown to 95% confluence on 35-mm-diameter plates and transfected with 4 μ g of each TGEV replicon, using 12 μ g of Lipofectamine 2000 (Invitrogen). At 36 hpt total intracellular RNA was extracted and used as a template for RT-PCR analysis of gene 7 mRNA following the procedure described above. Both, REP 1 and 2 were functional in all analyzed cells. However, in the case of REP 3 high levels of gene 7 mRNA were detected only in cells expressing TGEV N protein. These results indicated that N protein provided either in *cis* or in *trans* is required for efficient replicon activity. A quantitative analysis of gene 7 mRNA in BHK-pAPN cells or BHK-pAPN cells expressing TGEV N protein transfected with TGEV replicons was performed by real-time RT-PCR using the primers 7RS and LDVS, described above (Fig. 2C). A basal activity of REP 3 was detected, but the replicon activity increased more than 100-fold when N protein was provided in *cis* (REP 1 and REP 2 versus REP 3 activity in BHK-pAPN cells) and more than 1,000-fold when N protein was provided in *trans* (REP 3 activity in BHK-pAPN cells expressing TGEV N protein versus REP 3 activity in BHK-

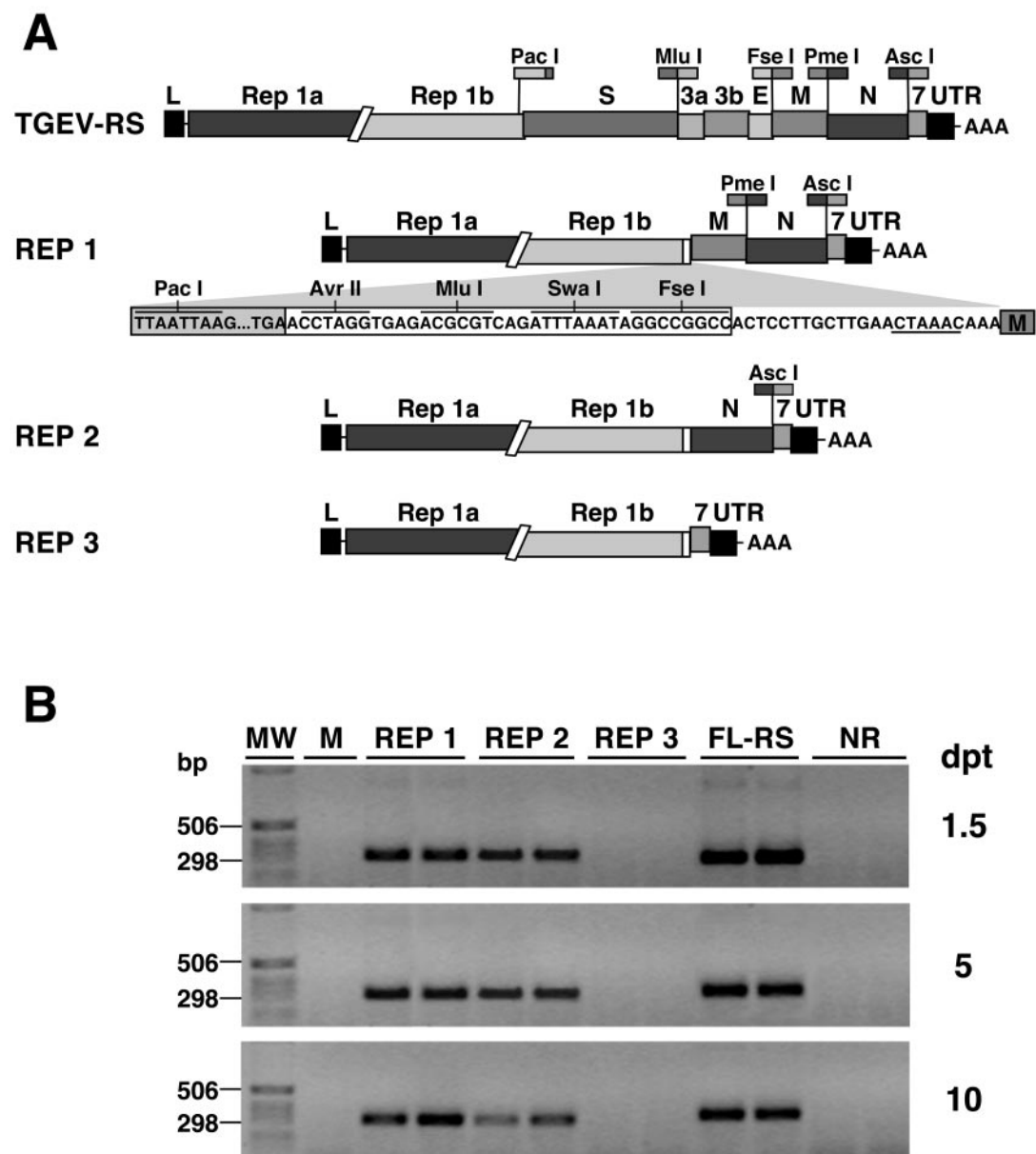


FIG. 1. Functional analysis of TGEV-derived replicons. (A) Genetic structures of TGEV-derived replicons. The genetic structures of the TGEV cDNA clone (TGEV-RS) and the replicons generated from this cDNA (REP 1, REP 2, and REP 3) are illustrated. To construct REP 1, a 125-bp fragment containing the restriction sites PacI, AvrII, MluI, SwaI, and FseI, generated by PCR using two overlapping oligonucleotides, was cloned in TGEV-RS digested with PacI and FseI. REP 2 and REP 3 were generated from REP 1 by deletion of fragments SwaI-PmeI and SwaI-AscI, respectively. To avoid possible interferences with the expression of heterologous genes, the TRS of the S gene, located at the 3' end of the replicase gene, was eliminated by introduction of three silent point mutations in its conserved core sequence. Letters and numbers above the bars indicate the viral genes. L, leader sequence; UTR, untranslated region. Relevant restriction sites are indicated. The core sequence is underlined. (B) Functional analysis of TGEV-derived replicons by RT-PCR. Human 293T cells were mock transfected (M) or transfected with TGEV replicons (REP 1, 2, and 3), the full-length cDNA clone (FL-RS), or a nonreplicative cDNA clone (NR), using Lipofectamine 2000 (Invitrogen) according to the manufacturer's specifications. Total intracellular RNA was isolated at 1.5, 5, and 10 dpt and analyzed by RT-PCR with specific oligonucleotides to detect gene 7 mRNA. Duplicate RT-PCR products amplified in parallel were resolved by electrophoresis in 1% agarose gels. MW, molecular weight markers.

pAPN cells). Interestingly, the activity of REP 1 and 2, carrying the N gene, was also increased when the N protein was provided in *trans*, suggesting that the N protein accumulation level is a critical factor or that the presence of N protein is important at the beginning of the replication process.

To discard the possibility that gene N mRNA itself was

responsible for TGEV replicon activity enhancement, the activity of REP 3 was analyzed by real-time RT-PCR in BHK-pAPN cells transformed with the Sindbis virus replicon expressing either gene N mRNA or the same mRNA mutated in order to eliminate N protein synthesis. To achieve this objective, three different mutants were constructed, in which one,

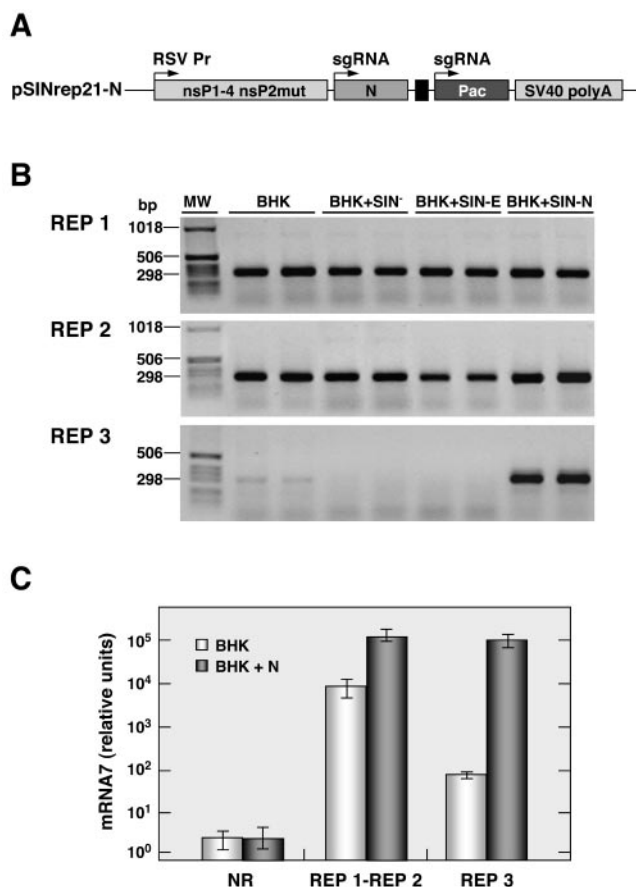


FIG. 2. Effect of N protein on TGEV-derived replicon activity. (A) Scheme of the Sindbis virus replicon construct encoding the TGEV N gene. RSV Pr, Rous sarcoma virus promoter; sgRNA, subgenomic RNA; nsP1-4, nonstructural proteins 1 to 4; nsP2mut, mutant of nonstructural protein 2; Pac, puromycin resistance gene; SV40 poly A, transcription-termination polyadenylation signal from simian virus 40. (B) RT-PCR analysis of gene 7 mRNA. BHK-pAPN cells (BHK) or BHK-pAPN cells transfected with either the Sindbis virus replicon pSINrep21 alone (BHK + SIN⁻) or pSINrep21 expressing TGEV E protein (BHK + SIN-E) or TGEV N protein (BHK + SIN-N) were transfected with TGEV replicons (REP 1, 2, and 3), using Lipofectamine 2000 (Invitrogen) following the manufacturer's specifications. At 36 hpt total intracellular RNA was extracted and used as the template for RT-PCR analysis with specific primers to detect gene 7 mRNA. Duplicate RT-PCR products amplified in parallel were resolved by electrophoresis in 1% agarose gels. MW, molecular weight markers. (C) Real-time RT-PCR quantification of gene 7 mRNA. The amount of mRNA7, expressed as relative units, was determined by real-time RT-PCR with specific oligonucleotides to detect gene 7 mRNA in RNA samples isolated at 36 hpt from BHK-pAPN cells (BHK) or BHK-pAPN cells expressing N protein (BHK + N) transfected with either a nonreplicative cDNA clone (NR), REP 1, REP 2, or REP 3. REP 1-REP 2 indicates mean values from REP 1 and 2, both encoding the N gene. Mean values from three experiments are represented, with standard deviations shown as error bars.

two, or three ATGs in the N gene coding sequence were mutated to ATC (Fig. 3A). Expression of gene 7 mRNA was used to study replicon activity following the procedure described above. The synthesis of gene N mRNA by the Sindbis virus replicon was analyzed by real-time RT-PCR using the antisense primer 5'-CCAATAACCAATCTGTTGATCCCT-3', complementary to nucleotides 187 to 219 of the N gene, and

the forward primer 5'-ACTTATGCTCGAGAGACTTTGTA CCC-3', spanning nucleotides 146 to 171 of the N gene. A clear reduction in the activity of REP 3 (more than 85%) was detected in BHK-pAPN cells expressing the different gene N mRNA mutants. However, no differences in the expression of gene N mRNA between the wild type and the different mutants were detected, indicating that N protein itself and not its mRNA was responsible for TGEV replicon enhancement (Fig. 3B).

These data indicated that N protein is not strictly essential, but it plays an important role as an enhancer of coronavirus replicon activity. The mechanism mediating N protein enhancement activity is unknown. N protein may be a critical component of the replication-transcription complex, or it may simply play a structural role to stabilize coronavirus replicon RNAs. This is the first time that a detailed and quantitative analysis of the requirement for N protein in coronavirus replication or transcription has been reported, although it was previously observed that N protein increases the rescue

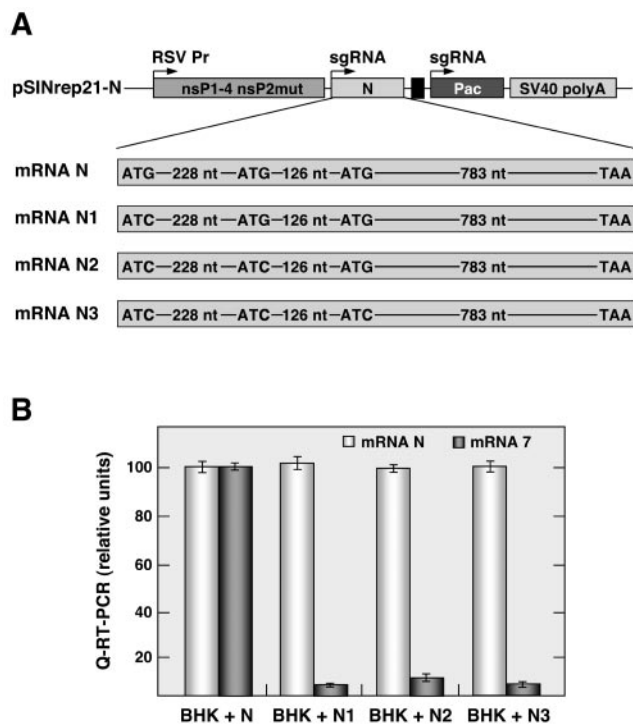


FIG. 3. Effect of N protein and its mRNA on TGEV-derived replicon activity. (A) Schematic representation of the Sindbis virus replicon encoding the different gene N mRNA mutants. The positions of the three ATGs in the N gene coding sequence, as well as the point mutations introduced, are indicated. The designations above the top bar are the same as those described for Fig. 2A. (B) Real-time RT-PCR quantification of genes 7 and N mRNAs. The amounts of genes 7 and N mRNAs, expressed as relative units, were determined by real-time RT-PCR with specific oligonucleotides in RNA samples isolated at 36 hpt from BHK cells expressing the wild-type gene N mRNA (BHK + N) or gene N mRNA with the first ATG mutated (BHK + N1), with the two first ATGs mutated (BHK + N2), or with the three ATGs mutated (BHK + N3) transfected with REP 3. Mean values from two experiments are represented, with standard deviations shown as error bars.

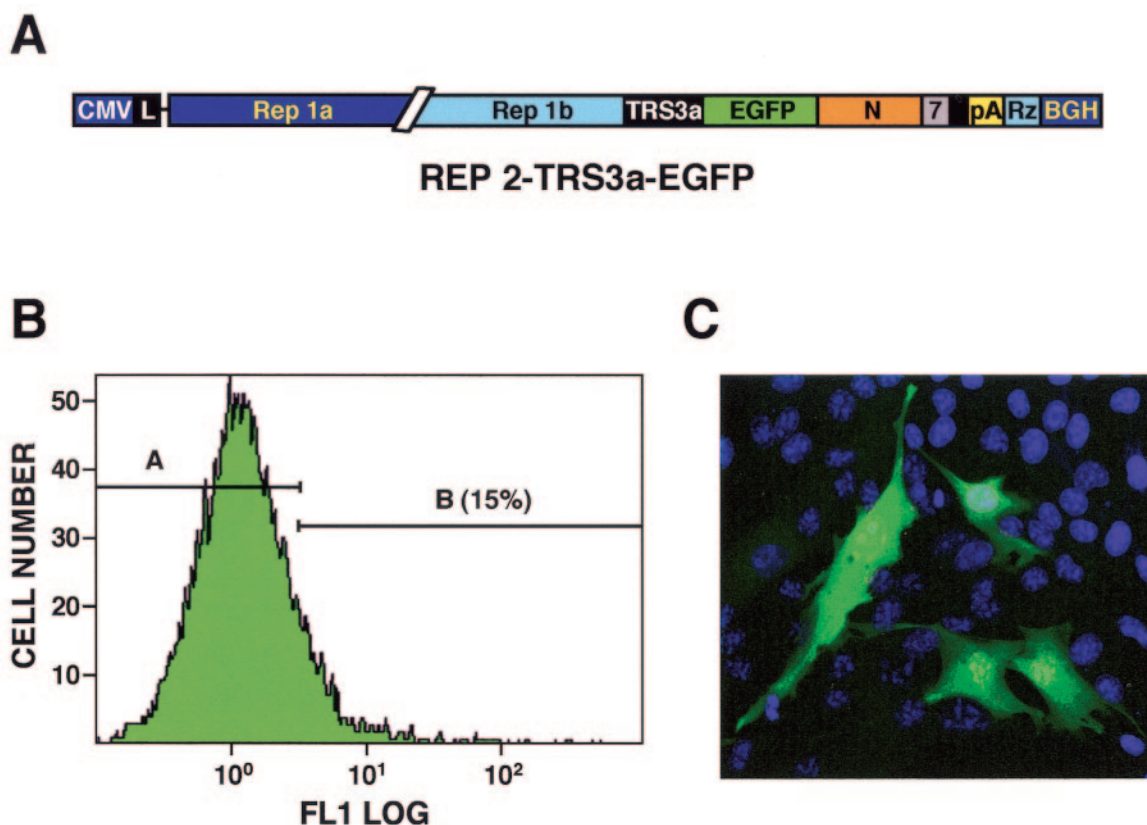


FIG. 4. Analysis of EGFP expression by TGEV-derived REP 2. (A) Schematic structure of the cDNA encoding the REP2-TRS3a-EGFP RNA with the EGFP gene under the control of gene 3a TRS. The numbers and letters inside the boxes indicate the viral genes. L, leader sequence; CMV, cytomegalovirus immediate-early promoter; pA, poly(A); Rz, hepatitis delta virus ribozyme; BGH, bovine growth hormone termination and polyadenylation sequences. (B) Flow cytometry analysis of EGFP expression in BHK-pAPN cells transfected with REP2-TRS3a-EGFP. A and B represent the EGFP-negative and -positive cell populations, respectively. Green fluorescence intensity is revealed in logarithmic units on the x axis. (C) EGFP expression analyzed by confocal microscopy of BHK-pAPN cells transfected with REP2-TRS3a-EGFP. Cell nuclei were stained with TOPRO 3 (Molecular Probes).

of coronavirus from infectious RNA transcripts (4, 32, 35, 37).

Heterologous gene expression from TGEV-derived replicons. In order to evaluate the potential of TGEV replicons to express heterologous genes, the expression of enhanced green fluorescent protein (EGFP) by REP 2 was analyzed in BHK-pAPN cells. A 1,115-bp cDNA fragment containing the EGFP gene downstream of native gene 3a TRS, comprising 262 nucleotides of the 5' TRS, the core sequence, and 23 nucleotides of the 3' TRS preceding the translation initiation codon, was amplified by PCR with the forward primer 5'-GCGGATATC TGTTTAAACGTGTGGCTACTAATAGGCTTAGTAG-3' (the restriction sites EcoRV and PmeI are underlined) and the reverse primer 5'-GGGGATATCGGCGCCAATATTTAA ATAAAATCACCATTGAG-3' (the restriction sites EcoRV and SfoI are underlined), using as the template the plasmid pBAC-TGEV-Δ3ab-EGFP (a full-length TGEV cDNA clone expressing EGFP under the control of TRS3a). The PCR product (TRS3a-EGFP) was digested with restriction endonuclease EcoRV and cloned into REP 2 digested with MluI and treated with Klenow enzyme to generate the plasmid REP2-TRS3a-EGFP (Fig. 4A). BHK-pAPN cells were grown to 95% confluence and transfected with REP2-TRS3a-EGFP, using

Lipofectamine 2000 (Invitrogen) as described in the previous section, and EGFP expression was analyzed at 36 hpt by flow cytometry (Fig. 4B) and confocal microscopy (Fig. 4C). A significant proportion of cells (18%) were transfected, and more than 80% of them expressed high levels of EGFP (Fig. 4C), indicating that a majority of cells transfected with the replicon expressed the heterologous gene. The strategy described here, in which the replicon encoding the N protein was launched from the cell nucleus by using the CMV promoter, allowed heterologous protein expression levels at least 1,000-fold higher than those previously reported with the human coronavirus 229E replicon (31).

The availability of functional replicons will significantly facilitate the study of the coronavirus replication mechanism and the production of safe TGEV-derived vectors for vaccination and possibly gene therapy.

We thank I. Sola and D. Escors for critically reading the manuscript and D. Dorado for technical assistance.

This work was supported by grants from the Ministerio de Ciencia y Tecnología (MCYT), Fort Dodge Veterinaria, and the European Community (Frame V, key action 2, Control of Infectious Disease Projects QLRT-2000-00876 and QLRT-FP6-5111060). C.G. received a predoctoral fellowship from the MCYT.

REFERENCES

- Agapov, E. V., I. Frolov, B. D. Lindenbach, B. M. Pragai, S. Schlesinger, and C. M. Rice. 1998. Noncytopathic Sindbis virus RNA vectors for heterologous gene expression. *Proc. Natl. Acad. Sci. USA* **95**:12989–12994.
- Almazán, F., J. M. González, Z. Pénzes, A. Izeta, E. Calvo, J. Plana-Durán, and L. Enjuanes. 2000. Engineering the largest RNA virus genome as an infectious bacterial artificial chromosome. *Proc. Natl. Acad. Sci. USA* **97**:5516–5521.
- Alonso, S., A. Izeta, I. Sola, and L. Enjuanes. 2002. Transcription regulatory sequences and mRNA expression levels in the coronavirus transmissible gastroenteritis virus. *J. Virol.* **76**:1293–1308.
- Casais, R., V. Thiel, S. G. Siddell, D. Cavanagh, and P. Britton. 2001. Reverse genetics system for the avian coronavirus infectious bronchitis virus. *J. Virol.* **75**:12359–12369.
- Delmas, B., J. Gelfi, H. Sjöström, O. Noren, and H. Laude. 1994. Further characterization of aminopeptidase-N as a receptor for coronaviruses. *Adv. Exp. Med. Biol.* **342**:293–298.
- Drosten, C., S. Gunther, W. Preiser, S. van der Werf, H. R. Brodt, S. Becker, H. Rabenau, M. Panning, L. Kolesnikova, R. A. Fouchier, A. Berger, A. M. Burguiere, J. Cinatl, M. Eickmann, N. Escriu, K. Grywna, S. Kramme, J. C. Manuguerra, S. Muller, V. Rickerts, M. Stürmer, S. Vieth, H. D. Klenk, A. D. Osterhaus, H. Schmitz, and H. W. Doerr. 2003. Identification of a novel coronavirus in patients with severe acute respiratory syndrome. *N. Engl. J. Med.* **348**:1967–1976.
- Enjuanes, L., W. Spaan, E. Snijder, and D. Cavanagh. 2000. Nidovirales, p. 827–834. *In* M. H. V. van Regenmortel, C. M. Fauquet, D. H. L. Bishop, E. B. Carsten, M. K. Estes, S. M. Lemon, M. A. Mayo, D. J. McGeoch, C. R. Pringle, and R. B. Wickner (ed.), *Virus taxonomy*. Academic Press, New York, N.Y.
- Fouchier, R. A., T. Kuiken, M. Schutten, G. van Amerongen, G. J. van Doornum, B. G. van den Hoogen, M. Peiris, W. Lim, K. Stohr, and A. D. Osterhaus. 2003. Aetiology: Koch's postulates fulfilled for SARS virus. *Nature* **423**:240.
- Frolov, I., T. A. Hoffman, B. M. Pragai, S. A. Dryga, H. V. Huang, S. Schlesinger, and C. M. Rice. 1996. Alphavirus-based expression vectors: strategies and applications. *Proc. Natl. Acad. Sci. USA* **93**:11371–11377.
- González, J. M., Z. Pénzes, F. Almazán, E. Calvo, and L. Enjuanes. 2002. Stabilization of a full-length infectious cDNA clone of transmissible gastroenteritis coronavirus by insertion of an intron. *J. Virol.* **76**:4655–4661.
- Herold, J., and R. Andino. 2001. Poliovirus RNA replication requires genome circularization through a protein-protein bridge. *Mol. Cell* **7**:581–591.
- Holmes, K. V., and L. Enjuanes. 2003. Virology. The SARS coronavirus: a postgenomic era. *Science* **300**:1377–1378.
- Izeta, A., C. Smerdou, S. Alonso, Z. Pénzes, A. Méndez, J. Plana-Durán, and L. Enjuanes. 1999. Replication and packaging of transmissible gastroenteritis coronavirus-derived synthetic minigenomes. *J. Virol.* **73**:1535–1545.
- Khromykh, A. A., and E. G. Westaway. 1997. Subgenomic replicons of the flavivirus Kunjin: construction and applications. *J. Virol.* **71**:1497–1505.
- Ksiazek, T. G., D. Erdman, C. S. Goldsmith, S. R. Zaki, T. Peret, S. Emery, S. Tong, C. Urbani, J. A. Comer, W. Lim, P. E. Rollin, S. F. Dowell, A. E. Ling, C. D. Humphrey, W. J. Shieh, J. Guarner, C. D. Paddock, P. Rota, B. Fields, J. DeRisi, J. Y. Yang, N. Cox, J. M. Hughes, J. W. LeDuc, W. J. Bellini, and L. J. Anderson. 2003. A novel coronavirus associated with severe acute respiratory syndrome. *N. Engl. J. Med.* **348**:1953–1966.
- Kuiken, T., R. A. Fouchier, M. Schutten, G. F. Rimmelzwaan, G. van Amerongen, D. van Riel, J. D. Laman, T. de Jong, G. van Doornum, W. Lim, A. E. Ling, P. K. Chan, J. S. Tam, M. C. Zambon, R. Gopal, C. Drosten, S. van der Werf, N. Escriu, J. C. Manuguerra, K. Stohr, J. S. Peiris, and A. D. Osterhaus. 2003. Newly discovered coronavirus as the primary cause of severe acute respiratory syndrome. *Lancet* **362**:263–270.
- Lohmann, V., F. Korner, J. Koch, U. Herian, L. Theilmann, and R. Bartenschlager. 1999. Replication of subgenomic hepatitis C virus RNAs in a hepatoma cell line. *Science* **285**:110–113.
- Marra, M. A., S. J. Jones, C. R. Astell, R. A. Holt, A. Brooks-Wilson, Y. S. Butterfield, J. Khattra, J. K. Asano, S. A. Barber, S. Y. Chan, A. Cloutier, S. M. Coughlin, D. Freeman, N. Girn, O. L. Griffith, S. R. Leach, M. Mayo, H. McDonald, S. B. Montgomery, P. K. Pandoh, A. S. Petrescu, A. G. Robertson, J. E. Schein, A. Siddiqui, D. E. Smailus, J. M. Stott, G. S. Yang, F. Plummer, A. Andonov, H. Artsob, N. Bastien, K. Bernard, T. F. Booth, D. Bowness, M. Czub, M. Drebot, L. Fernando, R. Flick, M. Garbutt, M. Gray, A. Grolla, S. Jones, H. Feldmann, A. Meyers, A. Kabani, Y. Li, S. Normand, U. Stroher, G. A. Tipples, S. Tyler, R. Vogrig, D. Ward, B. Watson, R. C. Brunham, M. Krajden, M. Petric, D. M. Skowronski, C. Upton, and R. L. Roper. 2003. The genome sequence of the SARS-associated coronavirus. *Science* **300**:1399–1404.
- Masters, P. S. 1999. Reverse genetics of the largest RNA viruses. *Adv. Virus Res.* **53**:245–264.
- Molenkamp, R., H. van Tol, B. C. Rozier, Y. van der Meer, W. J. Spaan, and E. J. Snijder. 2000. The arterivirus replicase is the only viral protein required for genome replication and subgenomic mRNA transcription. *J. Gen. Virol.* **81**:2491–2496.
- Ortego, J., D. Escors, H. Laude, and L. Enjuanes. 2002. Generation of a replication-competent, propagation-deficient virus vector based on the transmissible gastroenteritis coronavirus genome. *J. Virol.* **76**:11518–11529.
- Ortego, J., I. Sola, F. Almazán, J. E. Ceriani, C. Riquelme, M. Balasch, J. Plana, and L. Enjuanes. 2003. Transmissible gastroenteritis coronavirus gene 7 is not essential but influences *in vivo* virus replication and virulence. *Virology* **308**:13–22.
- Pénzes, Z., J. M. González, E. Calvo, A. Izeta, C. Smerdou, A. Méndez, C. M. Sánchez, I. Sola, F. Almazán, and L. Enjuanes. 2001. Complete genome sequence of transmissible gastroenteritis coronavirus PUR46-MAD clone and evolution of the purdue virus cluster. *Virus Genes* **23**:105–118.
- Repass, J. F., and S. Makino. 1998. Importance of the positive-strand RNA secondary structure of a murine coronavirus defective interfering RNA internal replication signal in positive-strand RNA synthesis. *J. Virol.* **72**:7926–7933.
- Rota, P. A., M. S. Oberste, S. S. Monroe, W. A. Nix, R. Campagnoli, J. P. Icenogle, S. Penaranda, B. Bankamp, K. Maher, M. H. Chen, S. Tong, A. Tamin, L. Lowe, M. Frace, J. L. DeRisi, Q. Chen, D. Wang, D. D. Erdman, T. C. Peret, C. Burns, T. G. Ksiazek, P. E. Rollin, A. Sanchez, S. Liffick, B. Holloway, J. Limor, K. McCaustland, M. Olsen-Rasmussen, R. Fouchier, S. Gunther, A. D. Osterhaus, C. Drosten, M. A. Pallansch, L. J. Anderson, and W. J. Bellini. 2003. Characterization of a novel coronavirus associated with severe acute respiratory syndrome. *Science* **300**:1394–1399.
- Sawicki, S. G., and D. L. Sawicki. 1998. A new model for coronavirus transcription. *Adv. Exp. Med. Biol.* **440**:215–219.
- Siddell, S., D. Sawicki, Y. Meyer, V. Thiel, and S. Sawicki. 2001. Identification of the mutations responsible for the phenotype of three MHV RNA-negative ts mutants. *Adv. Exp. Med. Biol.* **494**:453–458.
- Sola, I., S. Alonso, S. Zúñiga, M. Balasch, J. Plana-Durán, and L. Enjuanes. 2003. Engineering the transmissible gastroenteritis virus genome as an expression vector inducing lactogenic immunity. *J. Virol.* **77**:4357–4369.
- Stalcup, R. P., R. S. Baric, and J. L. Leibowitz. 1998. Genetic complementation among three panels of mouse hepatitis virus gene 1 mutants. *Virology* **241**:112–121.
- Thiel, V., J. Herold, B. Schelle, and S. G. Siddell. 2001. Infectious RNA transcribed *in vitro* from a cDNA copy of the human coronavirus genome cloned in vaccinia virus. *J. Gen. Virol.* **82**:1273–1281.
- Thiel, V., J. Herold, B. Schelle, and S. G. Siddell. 2001. Viral replicase gene products suffice for coronavirus discontinuous transcription. *J. Virol.* **75**:6676–6681.
- Thiel, V., N. Karl, B. Schelle, P. Disterer, I. Klage, and S. G. Siddell. 2003. Multigene RNA vector based on coronavirus transcription. *J. Virol.* **77**:9790–9798.
- Tijms, M. A., L. C. van Dinten, A. E. Gorbalenya, and E. J. Snijder. 2001. A zinc finger-containing papain-like protease couples subgenomic mRNA synthesis to genome translation in a positive-stranded RNA virus. *Proc. Natl. Acad. Sci. USA* **98**:1889–1894.
- Williams, G. D., R. Y. Chang, and D. A. Brian. 1999. A phylogenetically conserved hairpin-type 3' untranslated region pseudoknot functions in coronavirus RNA replication. *J. Virol.* **73**:8349–8355.
- Yount, B., K. M. Curtis, and R. S. Baric. 2000. Strategy for systematic assembly of large RNA and DNA genomes: transmissible gastroenteritis virus model. *J. Virol.* **74**:10600–10611.
- Yount, B., K. M. Curtis, E. A. Fritz, L. E. Hensley, P. B. Jahrling, E. Prentice, M. R. Denison, T. W. Geisbert, and R. S. Baric. 2003. Reverse genetics with a full-length infectious cDNA of severe acute respiratory syndrome coronavirus. *Proc. Natl. Acad. Sci. USA* **100**:12995–13000.
- Yount, B., M. R. Denison, S. R. Weiss, and R. S. Baric. 2002. Systematic assembly of a full-length infectious cDNA of mouse hepatitis virus strain A59. *J. Virol.* **76**:11065–11078.
- Ziebuhr, J., E. J. Snijder, and A. E. Gorbalenya. 2000. Virus-encoded proteinases and proteolytic processing in the Nidovirales. *J. Gen. Virol.* **81**:853–879.
- Zúñiga, S., I. Sola, S. Alonso, and L. Enjuanes. 2004. Sequence motifs involved in the regulation of discontinuous coronavirus subgenomic RNA synthesis. *J. Virol.* **78**:980–994.

A Point Mutation within the Replicase Gene Differentially Affects Coronavirus Genome versus Minigenome Replication

Carmen Galán, Luis Enjuanes,* and Fernando Almazán

Centro Nacional de Biotecnología, CSIC, Department of Molecular and Cell Biology, Campus Universidad Autónoma, Cantoblanco, Darwin St. 3, 28049 Madrid, Spain

Received 13 June 2005/Accepted 20 September 2005

During the construction of the *transmissible gastroenteritis virus* (TGEV) full-length cDNA clone, a point mutation at position 637 that was present in the defective minigenome DI-C was maintained as a genetic marker. Sequence analysis of the recovered viruses showed a reversion at this position to the original virus sequence. The effect of point mutations at nucleotide 637 was analyzed by reverse genetics using a TGEV full-length cDNA clone and cDNAs from TGEV-derived minigenomes. The replacement of nucleotide 637 of TGEV genome by a T, as in the DI-C sequence, or an A severely affected virus recovery from the cDNA, yielding mutant viruses with low titers and small plaques compared to those of the wild type. In contrast, T or A at position 637 was required for minigenome rescue in *trans* by the helper virus. No relationship between these observations and RNA secondary-structure predictions was found, indicating that mutations at nucleotide 637 most likely had an effect at the protein level. Nucleotide 637 occupies the second codon position at amino acid 108 of the pp1a polypeptide. This position is predicted to map in the N-terminal polypeptide papain-like proteinase (PLP-1) cleavage site at the p9/p87 junction. Replacement of G-637 by A, which causes a drastic amino acid change (Gly to Asp) at position 108, affected PLP-1-mediated cleavage in vitro. A correlation was found between predicted cleaving and noncleaving mutations and efficient virus rescue from cDNA and minigenome amplification, respectively.

Transmissible gastroenteritis virus (TGEV) is a member of the *Coronaviridae* family, which has recently been grouped with the *Arteriviridae* and the *Roniviridae* in the *Nidovirales* order (7, 14). Coronaviruses are the causative agents of a variety of diseases that impact animal and human health, including newly discovered members, such as the etiologic agent of severe acute respiratory syndrome and the human coronaviruses NL63 and HKU1 (13, 15, 32, 34, 42, 45).

Coronavirus genomes are single-stranded positive-sense infectious RNA molecules, the largest genomes known within RNA viruses, sharing a similar genome organization (4). The replicase gene encompasses the 5' two thirds of the genome and contains two overlapping open reading frames (ORF1a and ORF1b) that are translated by a ribosomal frameshift mechanism, leading to the expression of two large co-amino-terminal polypeptides, pp1a and pp1ab (5). Controlled proteolysis of these polypeptides by different virus-encoded proteinases regulates the release of the viral components of the replication-transcription complex (46, 47). This complex is believed to perform the synthesis of (i) the genome-size infectious RNA through a minus-strand antigenomic template, (ii) a nested set of subgenomic RNAs by a discontinuous transcription mechanism that most probably occurs during the synthesis of the negative strand (23, 36, 37), and (iii) deleted genomes, referred to as defective interfering RNAs or minigenomes, that conserve the *cis*-acting signals required for their amplification (18, 25, 28, 31, 44). Although RNA signals specifically involved in the synthesis of the subgenomic RNAs have been well characterized (40, 50), possible differences in

the machinery involved in the production of the different species of viral RNAs have not been addressed.

Coronavirus-derived minigenomes have been widely used to study different aspects of coronavirus replication. However, little is known about the precise mechanism and the selection pressure controlling their efficient maintenance in a helper-dependent system. Although findings in plant RNA viruses have illustrated some differences in the elements that regulate *trans*-accumulation of minigenomes from those controlling replication of the genomic RNA (33), no systematic studies of the elements controlling coronavirus replication have been reported in the context of both genomic and defective RNAs. The recent availability of coronavirus infectious clones and functional replicons allows reliable evaluation of the replication signals extrapolated from the analysis with defective RNAs.

Coronavirus ORF1a encodes one or two papain-like proteinases (PLPs) that are responsible for cleavage at one or two sites located in the N-proximal region of the replicase. The conservation and activity of the PLP proteinases vary among the three different coronavirus groups. PLP-1 plays a role in proteolytic processing at the relatively divergent amino terminus of the ORF1a polypeptide in coronavirus groups 1 and 2. In group 1 human coronavirus 229E (HCoV-229E), PLP-1 mediates the release of an amino-terminal 9-kDa protein and the adjacent 87-kDa protein, named p9 and p87, respectively (17, 48). The same processing pattern has been predicted by sequence comparison for the closely related TGEV coronavirus. For group 2 coronaviruses, including bovine coronavirus and mouse hepatitis virus (MHV), PLP-1 liberates the N-terminal p28 and p65 proteins of 28 and 65 kDa, respectively (2, 3, 10, 12). Functional explanations for these group-specific divergences have not yet been elucidated. Although PLP-1 cleavage products have been identified in virus-infected cells for MHV and HCoV-229E (17, 39), the specific activity of these proteins remains poorly characterized, and no

* Corresponding author. Mailing address: Department of Molecular and Cell Biology, Centro Nacional de Biotecnología, CSIC, Campus Universidad Autónoma, Cantoblanco, Darwin St. 3, 28049 Madrid, Spain. Phone: 34 91 585 4555. Fax: 34 91 585 4915. E-mail: L.Enjuanes@cnb.uam.es.

TABLE 1. Oligonucleotides used in this study

Oligonucleotide ^a	Method ^b	Oligonucleotide sequence, 5'→3' ^c	Amplicon size (bp)
Oli5'I Oli3'I	PCR 637 mut PCR 637 mut	AATTCGATGATAAGCTGTCAAC GGCACCTCTG(G/C/T)CAGTGCAGCAATTTTAAG	1,558
Oli5'D Oli3'D	PCR 637 mut PCR 637 mut	GCTCGCACTG(A/C/G)CAGAAGGTGCCATATATGTTG TCCTCTACTACTTTCCAAGCGTC	912
Oli5'I Oli3'D	PCR 637 mut PCR 637 mut	AATTCGATGATAAGCTGTCAAC TCCTCTACTACTTTCCAAGCGTC	2,449
Q-LBS7-VS Q-7-RS	Q-RT-PCR M33L Q-RT-PCR M33L	TACCAACACACCATGAAAAAATGC TCATTGACAGTTTTAAGATTGAAAGAGTG	123
Q-2060-VS Q-2167-RS	Q-RT-PCR DI-C Q-RT-PCR DI-C	TAAAGTTCCATCTTATGACATCGTTTATG TCTTGTGTGCCAGTTGGTATTTG	108
KpnI-M-E859-VS BamHI-STOP-S1315-RS	PCR PLP-1 PCR PLP-1	GCGCGGTACC ATG GAACATTTTTATCCTTATGG CGCGGGATCC TTA ACTTTGAATTTTCTCTTTTTC	1,375
KpnI-ORF1a-M1-VS BamHI-STOP-ORF1a-E610-RS	PCR ORF1a-S PCR ORF1a-S	GCGCGGTACC ATG AGTTCCAAACAATTCAAG CGCGGGATCC TAT TCAAATGATGAACCAAGTTTTC	1,830

^a VS, virus sense; RS, reverse sense.

^b PCR 637 mut, PCR for mutagenesis at position 637; Q-RT-PCR (M33L or DI-C), primers used for the quantification by Q-RT-PCR of the specified minigenome; PCR (PLP-1 or ORF1a), primers used for cloning of TGEV PLP-1 proteinase or substrate, respectively.

^c The mutated nucleotide is shown in parentheses. Restriction endonuclease sites used for cloning are underlined (KpnI, GGTACC; BamHI, GGATCC); translation initiation (ATG) and termination (TTA) codons are shown in bold.

counterparts of these proteins have been identified in other cellular or viral systems by sequence comparison. The MHV p28 and p65 polypeptides seem to be associated with intracellular membranes and colocalize with viral replication complexes (39). Additionally, it has been demonstrated that the expression of p28 induced cell cycle arrest in G₀/G₁ phase (8). Requirements of p28 and p65 for MHV-A59 replication have been studied by reverse genetics, showing that processing at the amino-terminal cleavage site of the replicase was not strictly essential for viral replication. However, predicted noncleaving mutations severely altered virus growth, protein processing, and RNA synthesis (11). No specific roles or functional requirements have been studied so far for the HCoV-229E or TGEV counterpart proteins p9 and p87.

In the present study, we characterize the effect of a point mutation at position 637 in the TGEV genome, affecting the predicted PLP-1 cleavage site at the p9/p87 junction. A mutational analysis at this position was performed by reverse genetics in the TGEV full-length cDNA clone (1) and in two TGEV-derived minigenomes of different sizes (20). Different nucleotides at position 637 were required for virus recovery from the cDNA and helper-mediated minigenome accumulation. A correlation was found between predicted cleaving and noncleaving mutations and the nucleotides selected for virus and minigenome replication, respectively.

MATERIALS AND METHODS

Cells and viruses. Baby hamster kidney cells stably transformed with the porcine aminopeptidase N gene (BHK-pAPN) (9) were grown in Dulbecco's modified Eagle's medium supplemented with 5% fetal calf serum and geneticin (1.5 mg/ml) as a selection agent. The TGEV PUR46-MAD strain (35) was used as the helper virus in the minigenome rescue system. Wild-type and recombinant viruses were grown and titrated in swine testis (ST) cells (27) as described previously (21).

Plasmid constructs and mutagenesis. To generate TGEV cDNAs with point mutations at genome position 637, the intermediate plasmid pBAC-TGEV (SrfI-NheI) (50) was used as a template for overlapping PCRs. A 1,558-bp PCR

fragment, containing the cytomegalovirus promoter and the first 647 nucleotides of the TGEV genome (30) (GenBank accession number AJ271965) with the desired point mutations at position 637, was amplified with oligonucleotides Oli5'I and Oli3'I, described in Table 1. A 912-bp overlapping PCR fragment including nucleotides 627 to 1531 of the TGEV genome with the corresponding mutations was amplified with primers Oli5'D and Oli3'D (Table 1). The final 2,449-bp PCR product was amplified using the overlapping PCR products and the outer oligonucleotides Oli5'I and Oli3'D, digested with SrfI and ApaLI, and cloned into the same restriction sites of plasmid pBAC-TGEV (SrfI-NheI), generating the plasmids pBAC-TGEV(SrfI-NheI)637A, pBAC-TGEV(SrfI-NheI)637T, pBAC-TGEV(SrfI-NheI)637C, and pBAC-TGEV(SrfI-NheI)637G. To introduce the engineered mutations in the TGEV infectious cDNA, the 5,277-bp SrfI-Clal fragment from pBAC-TGEV(SrfI-NheI) with the corresponding mutation was cloned into the same sites of pBAC-TGEV^{ΔClal} digested with the same restriction enzymes. Finally, the full-length cDNA clone was generated by cloning the ClaI⁴⁴¹⁷-ClaI⁹⁶¹⁵ fragment as previously described (1).

The construction of cDNAs encoding TGEV-derived RNA minigenome DI-C (9.7 kb) and DI-C-derived RNA minigenome M33 (3.3 kb) was previously described. These minigenomes were cloned under the T7 promoter and flanked at the 3' end by the hepatitis delta virus ribozyme (Rz) and the T7 terminator sequence (20). The M33L minigenome was constructed from the M33 minigenome by cloning a heterologous sequence of 16 bp, referred to as the linker, which does not interfere with the efficient rescue of this minigenome.

The original sequence of the DI-C and M33L cDNAs contains a T at position 637. M33L and DI-C cDNAs with point mutations at position 637 were generated by restriction fragment exchange. To get the M33L-637G cDNA, plasmid pBAC-TGEV-5'END, containing nucleotides 1 to 1496 from the TGEV genome, was digested with NaeI and SphI and the resulting 1,047-bp restriction fragment was cloned into the same sites of the plasmid coding for the M33L minigenome (pM33L). Plasmids pM33L-637A and pM33L-637C were generated by cloning the 646-bp NaeI-EcoRV restriction fragment from plasmids pBAC-TGEV(SrfI-NheI)637A and pBAC-TGEV(SrfI-NheI)637C, respectively, into plasmid pM33L digested with the same restriction enzymes.

DI-C cDNAs with point mutations at position 637 were generated by cloning the 1,229-bp NarI-EcoRV restriction fragment from pM33L with the corresponding mutations into the plasmid pDI-C (20) digested with the same restriction enzymes.

For the construction of plasmid pTGEV-PLP-1 encoding the PLP-1 domain, the genomic region corresponding to nucleotides 2889 to 4259, which encodes Glu859 to Ser1315 from pp1a, was amplified from pBAC-TGEV(SrfI-NheI) using primers KpnI-M-E859-VS and BamHI-STOP-S1315-RS, which introduced a KpnI site and a translation initiation codon at the 5' end and a translation

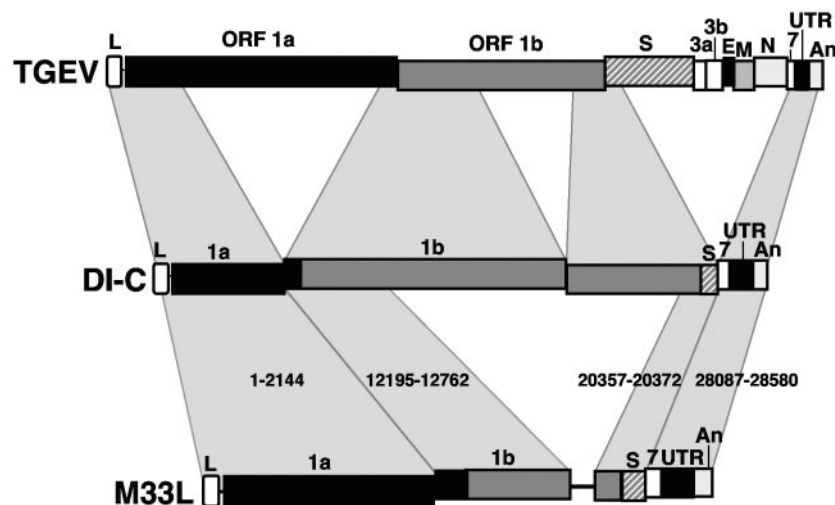


FIG. 1. Genetic structure of the TGEV genome and defective RNAs. The bar on top represents the TGEV virus genome, in which the different genes are depicted as boxes. The sequence relationship between the parental genome and minigenomes DI-C and M33L is indicated by shadowed polygons. The nucleotide positions of each discontinuous region in the TGEV genome are indicated for the minigenome M33L. L, leader; ORF1a and ORF1b, replicase genes; S, spike gene; 3a and 3b, nonstructural 3a and 3b genes; E, envelope protein gene; M, membrane protein gene; N, nucleoprotein gene; 7, nonstructural gene 7; UTR, untranslated region; An, poly(A).

termination codon followed by a BamHI site at the 3' end (Table 1). Following KpnI and BamHI digestion, a fragment containing the predicted PLP-1 domain was ligated into KpnI- and BamHI-cleaved pcDNA3 (Invitrogen).

For the generation of plasmids pTGEV-ORF1a(637G) and pTGEV-ORF1a(637A), the fragments encoding the N-terminal 610 amino acids from ORF1a, with either Gly or Asp at position 108, were obtained by PCR from plasmid pM33L-637G or pM33L-637A using primers KpnI-ORF1a-M1-VS and BamHI-STOP-ORF1a-E610-RS (Table 1) and cloned into KpnI- and BamHI-digested pcDNA3.

All cloning steps were checked by sequencing the PCR-amplified fragments and cloning junctions.

Transfection and recovery of recombinant viruses from cDNA clones. BHK-pAPN cells were grown to 95% confluence in 35-mm-diameter plates and transfected with 5 μ g of each pBAC-TGEV^{FL} plasmid containing the corresponding mutation at position 637 and 12 μ g of Lipofectamine 2000 (Invitrogen) in Optimum medium according to the manufacturer's specifications. Cells were incubated at 37°C for 6 h, then the transfection medium was discarded, and the cells were trypsinized and plated over a confluent monolayer of ST cells grown in a 35-mm-diameter plate. Cell supernatants were harvested for titration at 24, 48, and 72 h posttransfection.

In vitro transcription. Capped minigenome transcripts were generated from XhoI-linearized DNA templates using the T7 polymerase mMessage-mMachine kit (Ambion), according to the manufacturer's procedure. The length of the in vitro-transcribed RNAs was analyzed by denaturing 1% agarose-2.2 M formaldehyde gel electrophoresis.

Rescue of minigenomes by RNA transfection of helper virus-infected ST cells. ST cells were grown to 90% confluence and infected with TGEV PUR46-MAD at a multiplicity of infection of 10. At 2 h postinfection, cells were transfected with 1 μ g of in vitro-transcribed RNA using 3 μ g of Lipofectamine 2000 (Invitrogen) according to the manufacturer's specifications and incubated for 4 h at 37°C. Cells were then washed and incubated with fresh Dulbecco's modified Eagle's medium containing 2% fetal bovine serum. Supernatants from these cultures were used to infect fresh ST cell monolayers at 22 h postinfection and five serial passages were performed to amplify the minigenome RNA. Virus titer was determined in each passage as previously described (21). Total RNA was extracted in each passage by using the RNeasy Mini Kit (QIAGEN) according to the manufacturer's instructions. Additionally, total RNA was extracted at passage zero just after transfection from replicate samples, to further comparison of input minigenome RNA levels by real-time quantitative reverse transcription (RT)-PCR (Q-RT-PCR).

Q-RT-PCR. Analysis of minigenome rescue efficiency was performed by Q-RT-PCR on total RNA from each virus passage. cDNAs were synthesized at 50°C for 90 min with 5 units of Reflectase RT reverse transcriptase (Active Motif) using the antisense primers described in Table 1. The cDNAs generated were used as templates for real-time PCR amplification using specific primer

pairs designed with the Primer Express software (Table 1). The designed amplicons encompass specific discontinuous minigenome regions to discriminate hybridization with the viral genome. In the PCR step, SYBR Green master mix (Applied Biosystems) was used according to the manufacturer's specifications. Data acquisition was performed using an ABI PRISM 7000 sequence detection system (Applied Biosystems) and analyzed with ABI PRISM 7000 SDS version 1.0 software. All quantifications were performed in duplicate from two separate RNA transcripts of each minigenome which were transfected and rescued in parallel to validate the experimental data.

In vitro coupled transcription-translation and trans-cleavage assay. TGEV PLP-1 proteinase and ORF1a substrates cloned under the control of the T7 promoter were expressed using the TNT T7-coupled reticulocyte lysate system (Promega) according to the manufacturer's instructions. For the preparation of labeled substrates, in vitro transcription-translation reactions from pTGEV-ORF1a(637G) and pTGEV-ORF1a(637A) plasmids were performed in the presence of 20 μ Ci of L-[³⁵S]methionine (1,000 Ci/mmol at 10 mCi/ml; Amersham) and incubated at 30°C for 90 min. For PLP-1 translation, nonlabeled reactions were carried out in parallel in the presence of 1 mM methionine in accordance with the manufacturer's specifications, except that 100 μ M ZnCl₂ was added. Following completion of both substrate and enzyme reactions, cycloheximide was added to a final concentration of 0.6 μ g/ μ l to prevent further translation. Then, the substrate reaction mixtures were incubated overnight at 30°C with enzyme reaction mixture at enzyme-substrate (E:S) ratios of 1:2.8, 1:5, and 1:10 (vol/vol). Cleavage reactions were quenched with sodium dodecyl sulfate (SDS) gel load-

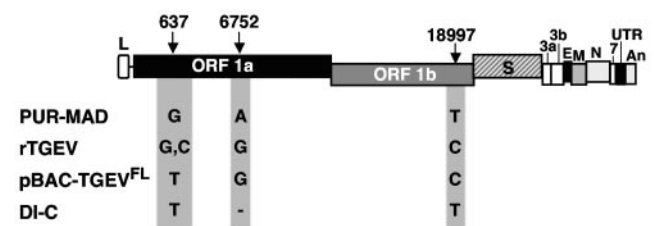


FIG. 2. Analysis of genetic markers of rTGEV viruses recovered from cDNA. The genetic structure of the TGEV cDNA clone and the positions of the genetic markers (indicated by arrows) are illustrated. Abbreviations are as in Fig. 1. The nucleotide sequence of the genetic markers is indicated for the parental virus (PUR-MAD), the rescued recombinant virus (rTGEV), the cDNA clone (pBAC-TGEV^{FL}), and the minigenome from which it was engineered (DI-C).

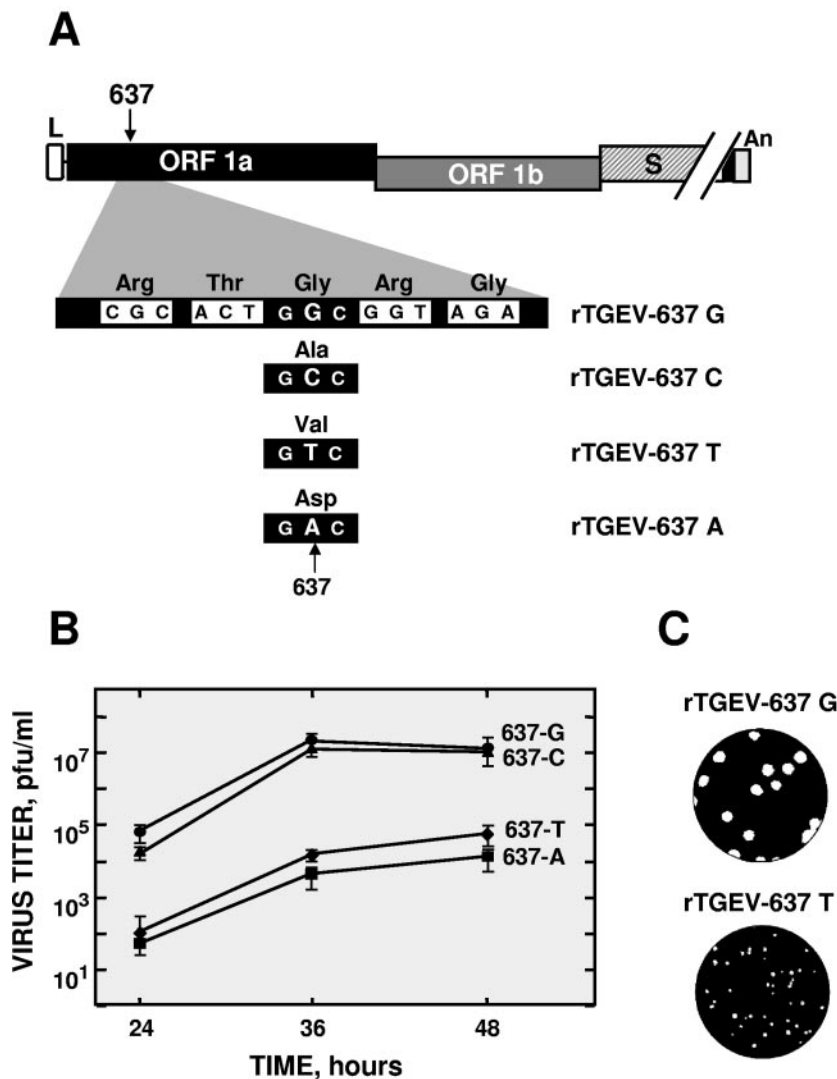


FIG. 3. Phenotype of recombinant viruses with point mutations at position 637. (A) Schematic diagram showing the genome organization of the TGEV cDNA constructs with the four nucleotide substitutions at position 637. Nucleotide substitutions and the resulting amino acid changes are indicated. The names of the recombinant viruses are indicated on the right side of the panel. (B) Virus rescue. TGEV cDNAs with the desired mutations at position 637 were transfected into BHK-pAPN cells as described in Materials and Methods and virus titers were determined at the indicated times posttransfection by plaque assay on ST cells. (C) Plaque morphology of the recombinant viruses. The phenotypes of rTGEV-637C and rTGEV-637A were similar to those of rTGEV-637G and rTGEV-637T, respectively.

ing buffer, and loaded onto SDS-10% polyacrylamide gels. After electrophoresis, gels were fixed in a methanol-acetic acid-water solution (5:1:4 [vol/vol/vol]) for 30 min, incubated in Amplify solution (Amersham NAMP100) for an additional 30 min, vacuum dried, and exposed to Kodak X-OMAT AR films at -80°C .

RESULTS

Reversion of a genetic marker in the recombinant TGEV at genome position 637. The TGEV cDNA clone pBAC-TGEV^{FL} was constructed from a defective minigenome (DI-C) derived from the respiratory strain PUR46-MAD (Fig. 1) (1). The pBAC-TGEV^{FL} construct contains two silent genetic markers at positions 6752 and 18997. A new cDNA clone containing one additional marker at position 637, which produces a conservative amino acid change (Gly to Val), was constructed (Fig. 2) because this nucleotide substitution was present in the

defective minigenome DI-C and in the DI-C-derived M33L minigenomes, which are stably and efficiently replicated by the helper virus (Fig. 1). Interestingly, when the genetic markers were sequenced in different recombinant TGEV (rTGEV) clones recovered from three independent cDNA transfections, all markers were maintained except nucleotide 637, which was replaced by the parental virus sequence G or by C, supporting the idea of the existence of a strong selective pressure on position 637, suggesting that either an RNA motif or a protein mapping within this domain had been affected and was playing an important role in the virus life cycle.

Mutations at genome position 637 severely affected virus recovery from the TGEV cDNA. To analyze the role of nucleotide 637 in virus replication, the influence of point mutations at this position on virus rescue was studied by mutational analysis

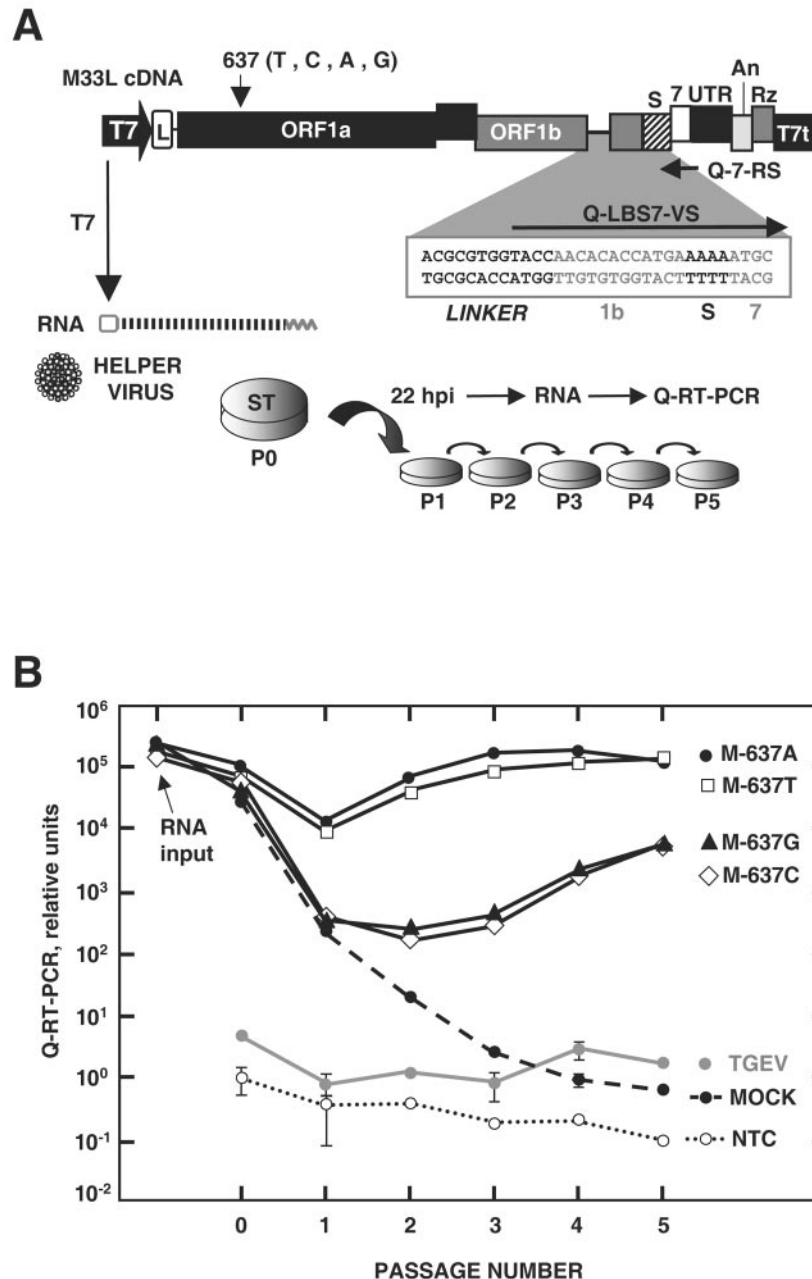


FIG. 4. Rescue of M33L mutants. (A) Scheme of the experimental procedure used to evaluate minigenome amplification. The bar on top represents the M33L cDNA, in which nucleotide 637 and the mutations introduced at this position are indicated. The discontinuous viral regions used for the specific design of the forward primer for Q-RT-PCR quantification are amplified under the M33L scheme. The specific oligonucleotides are indicated by arrows. T7-driven transcripts were transfected to helper TGEV-infected ST cells and amplified by five serial virus passages in confluent ST cell monolayers (P0 to P5). Total RNA was extracted from each virus passage to analyze minigenome amplification by Q-RT-PCR. Abbreviations are as in Fig. 1. T7, T7 promoter; Rz, ribozyme of hepatitis delta virus; T7t, T7 transcription termination signal. (B) Q-RT-PCR analysis. The RNA accumulation of M33L minigenome mutants at the successive viral passages, expressed in relative units, was determined by Q-RT-PCR. M33L-637 mutants are indicated at the right of each curve, named by the letter M followed by the corresponding mutation. Q-RT-PCR analyses of a noninfected (MOCK), a nontransfected (TGEV), and a nontemplated (NTC) control are also represented. RNA input indicates the quantification of a duplicate sample of each RNA transfection extracted at 4 h posttransfection. The standard deviation of replicate quantifications (when significant) is indicated by error bars.

using the infectious cDNA clone. Nucleotide substitution of the TGEV consensus sequence G at position 637 with C, T, or A produces an amino acid change at position 108 of the pp1a polypeptide from Gly to Ala, Val, or Asp, respectively (Fig. 3A). Recovered rTGEV viruses with the desired muta-

tions at this position were analyzed after transfection of the corresponding TGEV cDNAs. Virus rescue was analyzed by titration (Fig. 3B), showing a reduction of three logarithmic units for recombinant viruses with T or A at position 637, which correlated with smaller plaque morphology in relation to

that of the wild type (Fig. 3C). This phenotype was not stable and revertant viruses rapidly arose at early times posttransfection, resulting in fully revertant populations after the first viral passage. No differences in virus recovery efficiency or plaque size were observed between viruses containing C or G at this position. A correlation was found between mutations 637A and 637T, causing more drastic amino acid substitutions, and the recombinant viruses affected in replication, suggesting that the effect of these point mutations was most likely at the protein level.

Effect of mutations at position 637 on minigenome rescue efficiency. In contrast to the viral genome, TGEV-derived minigenomes that are stably and efficiently replicated by the helper virus contained a T at position 637 (Fig. 1) (20). To study the relevance of nucleotide substitutions at position 637 on minigenome replication, M33L minigenome mutants at this position were generated and their ability to be replicated and packaged by the PUR46-MAD helper virus was tested. ST cells were infected with TGEV PUR46-MAD at a multiplicity of infection of 10 and transfected with T7-driven transcripts of the M33L minigenome mutants. Supernatants from these cultures were passed five times into fresh ST cells in order to amplify the minigenomes (Fig. 4A). Total RNA was extracted at each passage and minigenome rescue efficiency was determined by Q-RT-PCR using specific primers (Fig. 4A and Table 1).

To determine if differences in RNA transfection efficiency could affect minigenome rescue, M33L minigenome transfections were performed in duplicate and total RNA was extracted from one duplicate sample at 4 h posttransfection and considered the RNA input. The amount of RNA input for each M33L mutant was quantified by Q-RT-PCR and no significant differences were detected (Fig. 4B). The relative rescue ability of M33L minigenomes was determined by Q-RT-PCR and was expressed as the relative amount of RNA accumulated at each viral passage (Fig. 4B). A control of cells transfected with the M33L minigenome but not infected with the helper virus was also quantified and considered the background level (Fig. 4B, MOCK). Nonspecific hybridization of the primers with the viral genome alone was monitored using an infected and mock-transfected control to assess the specificity of the assay (Fig. 4B, TGEV). The results revealed that only M33L minigenomes with A or T at position 637 were efficiently rescued. In contrast, as indicated above, nucleotide G or C at position 637 was required for efficient virus rescue from the TGEV infectious cDNA.

Interestingly, an increase in RNA accumulation was observed after the third viral passage for both M33L-637G and M33L-637C, suggesting a genotypic reversion at position 637. To examine this possibility, the entire M33L-637G and M33L-637C minigenomes were amplified from passage five by RT-PCR and sequenced. Both M33L-637G and M33L-637C minigenomes presented genotypic reversion, with the unique nucleotide substitution at position 637 from C or G to A. These results indicate that (i) basal levels of RNA amplification allowed the selection of revertants at position 637, (ii) this single nucleotide was responsible for the observed phenotypes, and (iii) T or A was clearly preferred at position 637 for minigenome rescue by the helper virus.

Additionally, to determine if differences in virus titers during the serial passages could be responsible for the differences observed in minigenome rescue, cell culture supernatants of

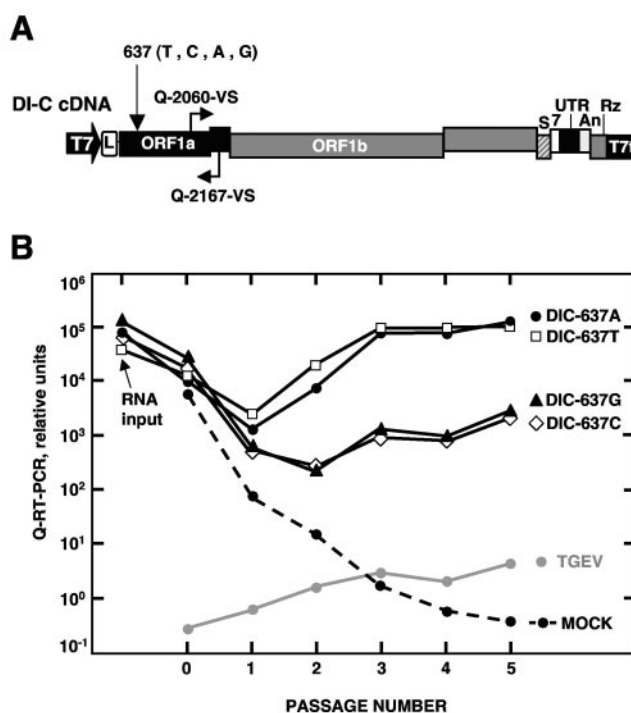


FIG. 5. Rescue of DI-C mutants. (A) Scheme of DI-C cDNA in which nucleotide 637 and the specific oligonucleotides designed to amplify the first discontinuous genomic region are indicated by arrows. Abbreviations are as in Fig. 4. (B) RNA quantification. The RNA accumulation of DI-C mutants at the successive viral passages, represented in relative units, was analyzed by Q-RT-PCR. Analyses of noninfected (MOCK) and nontransfected (TGEV) controls are also represented. RNA input indicates the quantification of a duplicate sample of each RNA transfection extracted at 4 h posttransfection. The standard deviation of replicate quantifications (when significant) is indicated by error bars.

each virus passage were titrated on ST cells. No differences in virus titers were observed that could explain the wide differences in minigenome amplification (data not shown).

Minigenome sequence requirements at position 637 are independent of minigenome size. The results shown above indicate that different sequence requirements at position 637 were necessary for efficient virus recovery from the infectious TGEV cDNA and for M33L minigenome rescue. Nevertheless, this observation could be a unique feature of the M33L minigenome or a general requirement for any TGEV-derived minigenome amplification. To extend our studies to a minigenome of a different size, the TGEV defective interfering RNA DI-C of 9.7 kb (Fig. 1) was used to perform a similar mutational analysis. ST cells were infected and transfected with the corresponding DI-C mutants and RNA accumulation was quantified by Q-RT-PCR for each virus passage using the same procedure and quantification controls as described for the M33L mutants (Fig. 5B). DI-C minigenomes containing T or A at position 637 were efficiently rescued, whereas DI-C mutants presenting G or C at the same position were severely affected in their ability to be rescued, showing the same behavior as the corresponding mutants of the M33L minigenome. On the basis of these results, it can be concluded that the same sequence requirements at nucleotide 637 are necessary for amplification

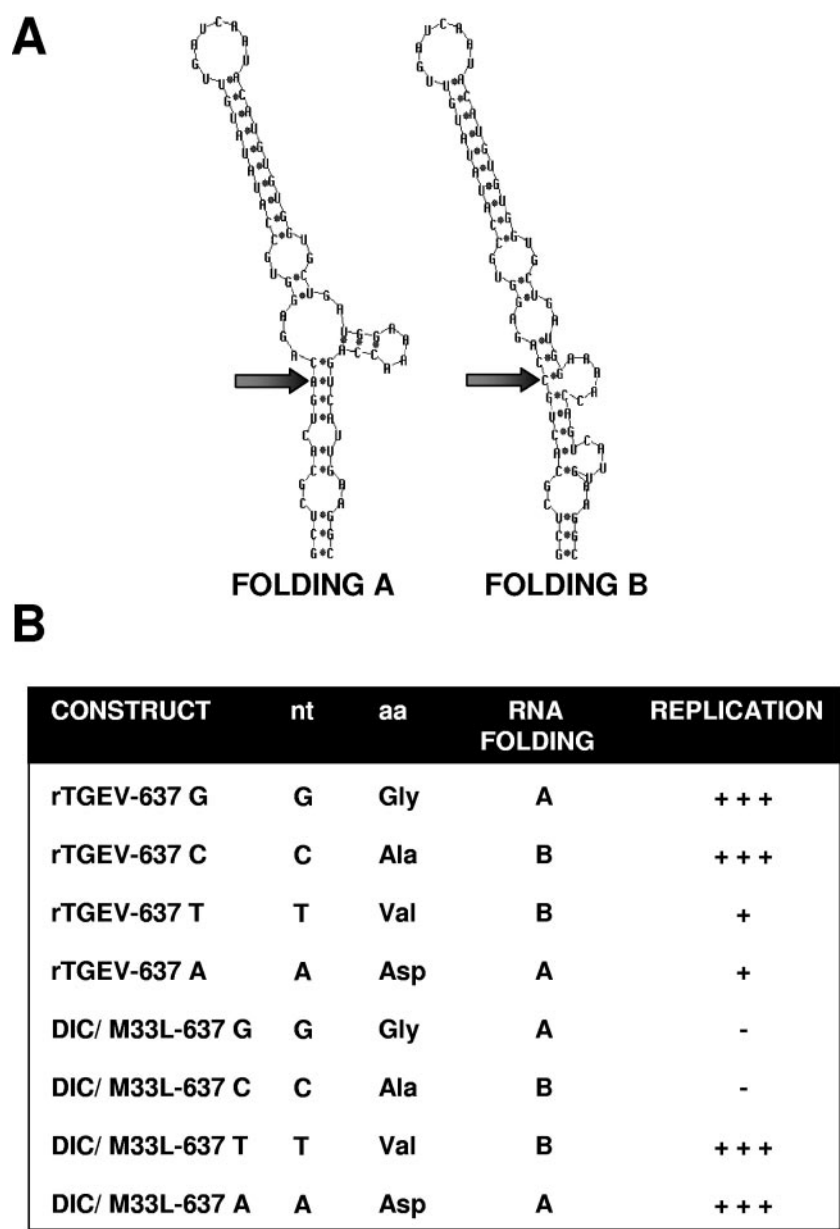


FIG. 6. In silico analysis of the RNA secondary structure of the TGEV domain containing nucleotide 637. (A) Secondary structure of the RNA sequence derived from the TGEV mutants. The secondary structure was predicted for each mutant using the Mfold 3.1 software. Predicted minimum energy structures were common for sequences containing G or A at position 637 (folding A) or those containing T or C (folding B). Only the RNA motif containing nucleotide 637 is shown, with an arrow indicating the nucleotide at this position. (B) Relationship between structure predictions of the RNA domain containing each substitution for nucleotide 637 and the observed phenotypes of the viruses and minigenomes with the corresponding mutations.

of both DI-C and M33L minigenomes regardless of their different sizes.

Effect of mutations at nucleotide 637 on RNA secondary-structure predictions. Consistent with previous results, it seems that both *cis* and *trans* TGEV replication was severely affected by single mutations at nucleotide 637, and different bases at this position were required and naturally selected for efficient virus or minigenome rescue. Mutations at nucleotide position 637 produce an amino acid change at position 108 of the pp1a and pp1ab polyproteins (Fig. 3A). Mutational analysis performed with the TGEV cDNA showed that mutations causing more

drastic amino acid changes correlated with those viruses that were more affected in replication. Furthermore, amino acid 108 mapped in the putative N-terminal PLP-1 cleavage site at the p9/p87 junction (Fig. 7A), as predicted by comparative sequence analysis with the human coronavirus HCoV-229E (16), suggesting that the effect of mutations at position 637 was at the protein level. Nevertheless, we cannot discard the idea that changes in the secondary structure of the RNA domain containing the nucleotide 637 could be responsible for the virus and minigenomes phenotypes.

To analyze this possibility, the relationship between the dif-

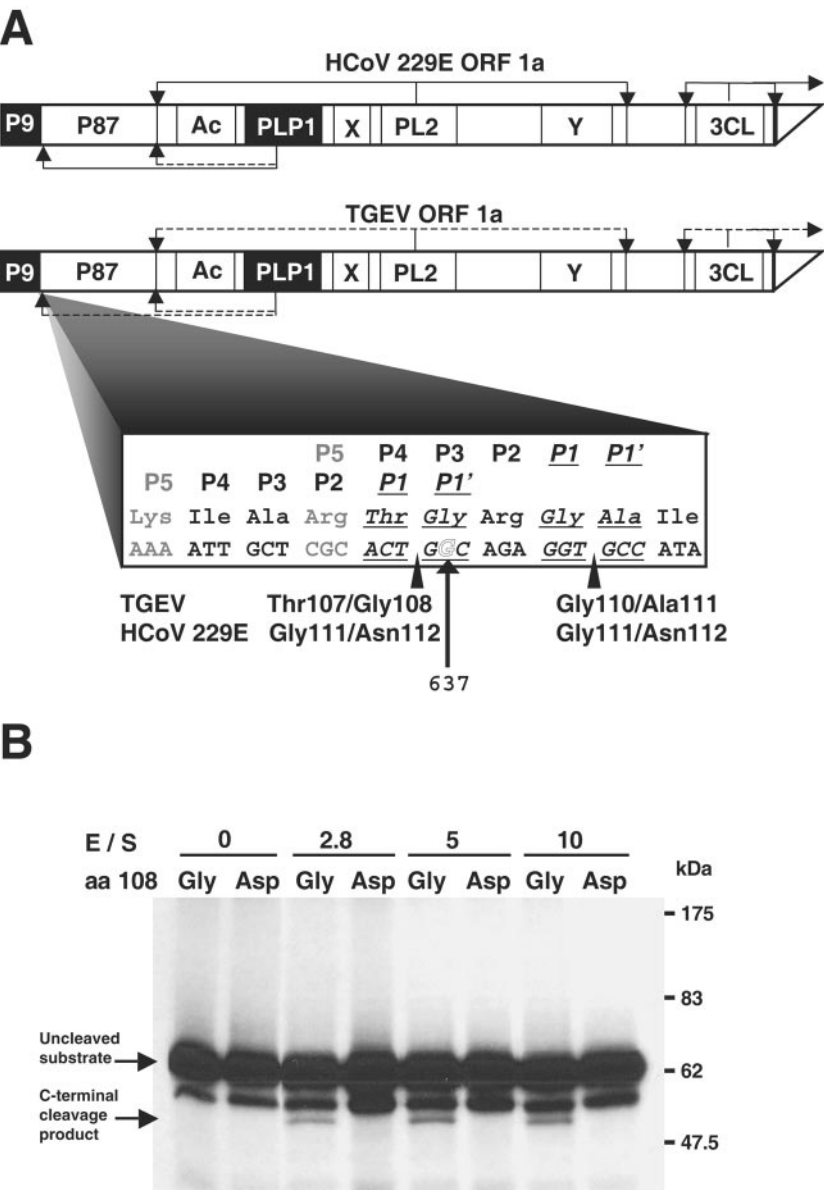


FIG. 7. Point mutations at nucleotide 637 affected N-terminal replicase processing in vitro. (A) Schematic representation of the amino-terminal region of ORF1a for HCoV-229E and TGEV. Arrows represent PLP-1 cleavage sites that have been experimentally shown (continuous lines) or inferred by sequence comparison (dotted lines). Replicase sequence containing the codon affected by mutations at position 637 is depicted. The two possible cleavage sites (black arrowheads) at the N terminus of the TGEV replicase were predicted by sequence alignment with HCoV-229E. Amino acids flanking the two potential PLP-1 cleavage sites are designated according to the nomenclature introduced by Schechter and Berger (38) as NH₂-P3-P2-P1 ↓ P1'-P2'-P3'-COOH. (B) *trans*-Cleavage assay. The PLP-1 domain and the pp1a N-terminal 610 amino acids with the wild-type sequence G at position 637 or an A at the same position, leading to the least conservative amino acid change (Gly108Asp), were generated as described in Materials and Methods. After translation, 1 volume of [³⁵S]methionine-labeled substrate was incubated for 14 h at 30°C in the absence or presence of 2.8, 5, or 10 volumes of enzyme reaction mixture. The cleavage reaction products were separated by sodium dodecyl sulfate–10% polyacrylamide gel electrophoresis and analyzed by fluorography. Molecular mass markers are given on the left in kDa, and the uncleaved substrate and C-terminal cleavage product are indicated by arrows on the right.

ferent mutations at position 637, the replication ability of the corresponding virus and minigenome mutants, and the secondary structure of the RNA domain containing this nucleotide were studied (Fig. 6B). Structure predictions were performed using the Mfold software (26, 49) and constructs derived from the 5' end of the TGEV genome with point mutations at position 637 were analyzed. Two folding patterns were obtained, one common for constructs containing G or A at po-

sition 637 (Fig. 6A, folding A) and another one for those containing C or T at the same position (Fig. 6A, folding B). A more detailed analysis considering sequences of different lengths and both positive and negative RNA strand polarities was consistent with this observation (data not shown). As described above, different phenotypes were observed for virus or minigenomes containing G or A and for those presenting T or C. Therefore, no correlation was found between predictions of

RNA folding and virus or minigenome replication ability, supporting the idea that the effect of mutations at position 637 was most probably at the protein level.

Mutations at nucleotide 637 affected N-terminal replicase processing in vitro by PLP-1. The results reported so far point toward an effect of the mutations at position 637, associated with processing at the N-terminal region of the TGEV replicase. The PLP-1 cleavage site can occupy a variable position at the amino terminus of the replicase polyproteins, as has been reported for different coronaviruses (17, 19). The PLP-1 of HCoV-229E, closely related to TGEV coronavirus, is responsible for the release of the N-terminal p9 protein from the pp1a and pp1ab polyproteins, with cleavage occurring between residues Gly111 and Asn112 (Fig. 7A). No experimental data have been reported for TGEV polyprotein processing, although potential cleavage sites are proposed by comparative sequence analysis with HCoV-229E.

Amino acid 108, affected by mutations at nucleotide position 637, occupies the P1' position at the Thr107/Gly108 theoretical cleavage site or the P3 position at the Gly110/Ala111 alternative cleavage site (Fig. 7A) (17). In both cases, a drastic amino acid change could probably affect the processing mediated by the PLP-1, as described in mutagenesis studies and sequence analyses of three PLP-1 cleavage sites (3, 12, 19). To support this hypothesis, a *trans*-cleavage assay was performed with the putative TGEV PLP-1 domain and an N-terminal pp1a substrate with either the wild-type sequence G at position 637 or an A at the same position, leading to the least conservative amino acid change (Gly108Asp). A fragment containing the predicted PLP-1 domain was cloned to undergo T7-driven transcription coupled to translation. An in vitro translation product representing the amino-terminal 610 amino acids of TGEV pp1a was used as the ³⁵S-labeled substrate in the *trans*-cleavage assay. This sequence was selected because it is the one encoded by the first 2,144 nucleotides present at the 5' continuous genomic region in both M33L and DI-C minigenomes, which are efficiently rescued by the helper virus.

After translation was accomplished, different enzyme-substrate volume ratios were tested for cleavage by combining translation mixtures including the substrate and the PLP-1 (Fig. 7B). In the absence of enzyme, only the labeled substrate with a molecular mass of 67 kDa was detected, together with a minor protein band that was probably the result of premature termination events during transcription or translation. When PLP-1 was included, a processed form of the substrate was detected with the substrate containing a Gly residue at position 108, but not with that presenting an Asp residue at the same position. The PLP-1 cleavage efficiency observed was similar for the different enzyme-substrate ratios tested. The pp1a proteolytic product correlated with the expected size of the C-terminal cleavage product generated after the release of the N-terminal p9 protein, and was not the result of a spontaneous degradation event from this particular substrate, since this processed protein was not detected when both substrates were incubated in parallel without enzyme. In this experiment, no N-terminal p9 cleavage product was detected, probably due to its reduced labeling (the cleaved peptide contains only two radiolabeled residues, compared to seven present in the C-terminal cleavage product) and to a reduced efficiency of PLP-1. These results indicated that

mutations at nucleotide 637 affected the in vitro cleavage of the TGEV replicase mediated by PLP-1.

DISCUSSION

The effect of point mutations at position 637, mapping at the putative N-terminal cleavage site of the TGEV replicase, on virus and minigenome replication has been characterized. Mutations at nucleotide 637 affected the in vitro cleavage of the TGEV replicase mediated by PLP-1. Different mutations at position 637 were required for virus recovery from the cDNA and minigenome accumulation. A correlation has been established between those mutations most likely interfering with proper N-terminal replicase processing and the limited rescue efficiency of the recombinant viruses.

Despite the relative flexibility of the TGEV infectious clone for genetic modification (29), a point mutation that was present in the defective interfering minigenome DI-C, which is stably and efficiently amplified in *trans* by the helper virus, was rapidly reverted in the genome of the recombinant progeny virions, indicating a strong selective pressure at this position.

The correlation between nucleotide 637 and the N-terminal cleavage of the replicase by the viral proteinase PLP-1 was based on previous sequence alignments reported for the HCoV-229E and TGEV replicase polyproteins (17). HCoV-229E PLP-1 activity and substrate specificity has been reported (16, 17). According to these studies, TGEV PLP-1 mediates N-terminal cleavage at either Thr107/Gly108 or Gly110/Ala111. Nucleotide substitutions at position 637 affect the nature of amino acid 108, which occupies the P1' and P3 residues relative to the Thr107/Gly108 and Gly110/Ala111 cleavage sites, respectively. In this report, it has been shown that a nonconservative amino acid change at this position affected the cleavage mediated by PLP-1 using an in vitro *trans*-cleavage assay. The processing of PLP-1 was limited, as previously described for in vitro assays with minimal domains of coronavirus PLP proteinases acting in *trans* (16, 41). These results could be due to the relatively reduced length of the substrate, to the lack of flanking domains required for efficient processing in the PLP-1 construct tested (47), or both.

The effect of mutations at position 637 was studied in two systems, the full-length TGEV genome and TGEV-derived defective minigenomes. The introduction of mutations predictably affecting proper replicase processing in the full-length TGEV genome caused a dramatic reduction in the virus recovery and a small-plaque-size phenotype, suggesting a critical role in virus replication of the proteins related to this cleavage site. PLP-1 cleavage products have been identified in virus-infected cells for HCoV-229E coronavirus, such as an early-released p9 polypeptide, a p87 adjacent protein, and several polypeptides with apparent molecular masses of 93 kDa (p93), 170 kDa (p170), and 230 kDa (p230), whose identity and potential precursor-product relationships remain to be determined (16). However, no functional data have been reported so far for these proteins and there is no obvious sequence comparison from which a putative function could be deduced.

The equivalent proteins of MHV replicase are generated from a relatively divergent amino-terminal domain. Thus, the p28 and p65 polypeptides are the counterparts of the p9 and p87 proteins of HCoV-229E, respectively. Biochemical studies have shown that the MHV p28 and p65 polypeptides seem to

be associated with intracellular membranes and colocalize with viral replication complexes (6, 39). In addition, recent data correlated p28 expression with cell cycle arrest in G₀/G₁ phase (8). Our data correlated with the mutational analysis performed in the MHV-A59 infectious clone at the N-terminal PLP-1 cleavage site of the MHV replicase, in which predicted noncleaving mutations severely affected virus growth, protein processing, and RNA synthesis (11). However, in contrast to the MHV mutants, the TGEV mutants characterized in our work harboring predicted noncleaving mutations were not stable and could not be characterized further, since rapidly reverted to the wild-type sequence. Although group-specific divergences could explain different requirements at this cleavage site, it seems clear that the abrogation of this specific processing event severely compromises the virus viability for both the MHV and TGEV coronaviruses.

In contrast, TGEV-derived minigenomes with predicted noncleaving mutations at the N-terminal cleavage site were efficiently rescued by the helper virus, whereas predicted cleaving mutations negatively affected their ability to be rescued in *trans*. The observed differences might have relevant implications. (i) Sequence requirements for minigenome amplification in *trans* could differ from those required for genome replication in *cis*, reinforcing the need to generate replicons to study coronavirus replication instead of using defective minigenomes. (ii) Different mechanisms seem to operate in the replication of full-length viral genomes and minigenomes, allowing the minigenomes to take advantage of the replication machinery of the helper virus to modulate viral infectivity. (iii) The defective minigenomes M33L and DI-C encode a replicase-derived fusion protein that could interfere with minigenome amplification due to the different mature products generated after processing. This fusion protein comprises the first 610 amino acids from the replicase N terminus, merged in phase with the last 58 amino acids of pp1a, and the slippery sequence that could produce the translation frameshift resulting in an extended version of this fusion protein.

Although there is no consensus about the translation requirements of the defective coronavirus RNAs characterized so far, it seems that translation, not the encoded sequence, might confer an advantage for the efficient propagation of these RNAs, most probably by coupling the translation and replication mechanisms (22, 24, 43). The putative cleavage of this fusion protein in the minigenomes studied, M33L and DI-C, by the helper virus PLP-1 proteinase activity would release the amino-terminal p9 polypeptide and a fusion protein containing a truncated form of the p87 polypeptide. A plausible interpretation is that the truncated polypeptide might interfere with minigenome amplification by sequestering the minigenome in nonfunctional replication complexes, interfering with the access of the minigenome to the helper virus replication machinery. Further experiments will be required to characterize the strategy developed by minigenomes to undergo effective *trans*-amplification and the specific function of each component of coronavirus replicase polypeptides in the synthesis of the different species of viral RNA.

ACKNOWLEDGMENTS

We thank S. Zúñiga for critically reading the manuscript and helpful discussions.

This work was supported by grants from the Ministerio de Educación y Ciencia (MEC) of Spain, Fort Dodge Veterinaria, and the European Community (Frame V, Key Action 2, Control of Infectious Disease Projects QLRT-2000-00874, and Frame VI, SP22-CT-2004-511060). C.G. received a predoctoral fellowship from the MEC.

REFERENCES

- Almazán, F., J. M. González, Z. Péntes, A. Izeta, E. Calvo, J. Plana-Durán, and L. Enjuanes. 2000. Engineering the largest RNA virus genome as an infectious bacterial artificial chromosome. *Proc. Natl. Acad. Sci. USA* **97**: 5516–5521.
- Baker, S. C., C. K. Shieh, L. H. Soe, M. F. Chang, D. M. Vannier, and M. M. C. Lai. 1989. Identification of a domain required for autoproteolytic cleavage of murine coronavirus gene A polyprotein. *J. Virol.* **63**:3693–3699.
- Bonilla, P. J., S. A. Hughes, and S. R. Weiss. 1997. Characterization of a second cleavage site and demonstration of activity in *trans* by the papain-like proteinase of the murine coronavirus mouse hepatitis virus strain A59. *J. Virol.* **71**:900–909.
- Brian, D. A., and R. S. Baric. 2005. Coronavirus genome structure and replication. *Curr. Top. Microbiol. Immunol.* **287**:1–30.
- Brierley, I., P. Digard, and S. C. Inglis. 1989. Characterization of an efficient coronavirus ribosomal frameshifting signal: requirement for an RNA pseudoknot. *Cell* **57**:537–547.
- Brockway, S. M., X. T. Lu, T. R. Peters, T. S. Dermody, and M. R. Denison. 2004. Intracellular localization and protein interactions of the gene 1 protein p28 during mouse hepatitis virus replication. *J. Virol.* **78**:11551–11562.
- Cavanagh, D., D. A. Brian, P. Britton, L. Enjuanes, M. C. Horzinek, M. M. C. Lai, H. Laude, P. G. W. Plagemann, S. Siddell, W. Spaan, and P. J. Talbot. 1997. *Nidovirales*: a new order comprising *Coronaviridae* and *Arteriviridae*. *Arch. Virol.* **142**:629–635.
- Chen, C. J., K. Sugiyama, H. Kubo, C. Huang, and S. Makino. 2004. Murine coronavirus nonstructural protein p28 arrests cell cycle in G₀/G₁ phase. *J. Virol.* **78**:10410–10419.
- Delmas, B., J. Gelfi, H. Sjöström, O. Noren, and H. Laude. 1993. Further characterization of aminopeptidase-N as a receptor for coronaviruses. *Adv. Exp. Med. Biol.* **342**:293–298.
- Denison, M. R., S. A. Hughes, and S. R. Weiss. 1995. Identification and characterization of a 65-kDa protein processed from the gene 1 polyprotein of the murine coronavirus MHV-A59. *Virology* **207**:316–320.
- Denison, M. R., B. Yount, S. M. Brockway, R. L. Graham, A. C. Sims, X. Lu, and R. S. Baric. 2004. Cleavage between replicase proteins p28 and p65 of mouse hepatitis virus is not required for virus replication. *J. Virol.* **78**:5957–5965.
- Dong, S., and S. Baker. 1994. Determinants of the p28 cleavage site recognized by the first papain-like cysteine proteinase of murine coronavirus. *Virology* **204**:541–549.
- Drosten, C., S. Günther, W. Preiser, S. van der Werf, H. R. Brodt, S. Becker, H. Rabenau, M. Panning, L. Kolesnikova, R. A. M. Fouchier, A. Berger, A. M. Burguiere, J. Cinatl, M. Eickmann, N. Escriou, K. Grywna, S. Kramme, J. C. Manuguerra, S. Muller, W. Rickerts, M. V. Stürmer, S., H. D. Klenk, and A. D. M. E. Osterhaus. 2003. Identification of a novel coronavirus in patients with severe acute respiratory syndrome. *N. Engl. J. Med.* **348**: 1967–1976.
- Enjuanes, L., W. Spaan, E. Snijder, and D. Cavanagh. 2000. *Nidovirales*, p. 827–834. In R. B. Wickner (ed.), *Virus taxonomy: classification and nomenclature of viruses*. Academic Press, San Diego, Calif.
- Esper, F., C. Weibel, D. Ferguson, M. L. Landry, and J. S. Kahn. 2005. Evidence of a novel human coronavirus that is associated with respiratory tract disease in infants and young children. *J. Infect. Dis.* **191**:492–498.
- Herold, J., A. E. Gorbalenya, V. Thiel, B. Schelle, and S. G. Siddell. 1998. Proteolytic processing at the amino terminus of human coronavirus 229E gene 1-encoded polyproteins: identification of a papain-like proteinase and its substrate. *J. Virol.* **72**:910–918.
- Herold, J., V. Thiel, and S. G. Siddell. 1998. Characterization of a papain-like cysteine-proteinase encoded by gene 1 of the human coronavirus HCV 229E. *Adv. Exp. Med. Biol.* **440**:141–148.
- Hofmann, M. A., P. B. Sethna, and D. A. Brian. 1990. Bovine coronavirus mRNA replication continues throughout persistent infection in cell culture. *J. Virol.* **64**:4108–4114.
- Hughes, S. A., P. J. Bonilla, and S. R. Weiss. 1995. Identification of the murine coronavirus p28 cleavage site. *J. Virol.* **69**:809–813.
- Izeta, A., C. Smerdou, S. Alonso, Z. Péntes, A. Méndez, J. Plana-Durán, and L. Enjuanes. 1999. Replication and packaging of transmissible gastroenteritis coronavirus-derived synthetic minigenomes. *J. Virol.* **73**:1535–1545.
- Jiménez, G., I. Correa, M. P. Melgosa, M. J. Bullido, and L. Enjuanes. 1986. Critical epitopes in transmissible gastroenteritis virus neutralization. *J. Virol.* **60**:131–139.
- Kim, Y. N., M. M. C. Lai, and S. Makino. 1993. Generation and selection of coronavirus defective interfering RNA with large open reading frame by RNA recombination and possible editing. *Virology* **194**:244–253.
- Lai, M. M. C., and D. Cavanagh. 1997. The molecular biology of coronaviruses. *Adv. Virus Res.* **48**:1–100.

24. Liao, C. L., and M. M. Lai. 1995. A cis-acting viral protein is not required for the replication of a coronavirus defective-interfering RNA. *Virology* **209**: 428–436.
25. Makino, S., F. Taguchi, and K. Fujiwara. 1984. Defective interfering particles of mouse hepatitis virus. *Virology* **133**:9–17.
26. Mathews, D. H., J. Sabina, M. Zuker, and D. H. Turner. 1999. Expanded sequence dependence of thermodynamic parameters improves prediction of RNA secondary structure. *J. Mol. Biol.* **288**:911–940.
27. McClurkin, A. W., and J. O. Norman. 1966. Studies on transmissible gastroenteritis of swine. II. Selected characteristics of a cytopathogenic virus common to five isolates from transmissible gastroenteritis. *Can. J. Comp. Med. Vet. Sci.* **30**:190–198.
28. Méndez, A., C. Smerdou, A. Izeta, F. Gebauer, and L. Enjuanes. 1996. Molecular characterization of transmissible gastroenteritis coronavirus defective interfering genomes: packaging and heterogeneity. *Virology* **217**:495–507.
29. Ortega, J., I. Sola, F. Almazan, J. E. Ceriani, C. Riquelme, M. Balasch, J. Plana-Durán, and L. Enjuanes. 2003. Transmissible gastroenteritis coronavirus gene 7 is not essential but influences *in vivo* virus replication and virulence. *Virology* **308**:13–22.
30. Penzes, Z., J. M. González, E. Calvo, A. Izeta, C. Smerdou, A. Mendez, C. M. Sánchez, I. Sola, F. Almazán, and L. Enjuanes. 2001. Complete genome sequence of transmissible gastroenteritis coronavirus PUR46-MAD clone and evolution of the Purdue virus cluster. *Virus Genes* **23**:105–118.
31. Penzes, Z., K. Tibbles, K. Shaw, P. Britton, T. D. K. Brown, and D. Cavanagh. 1994. Characterization of a replicating and packaged defective RNA of avian coronavirus infectious bronchitis virus. *Virology* **203**:286–293.
32. Perlman, S., T. E. Lane, and M. J. Buchmeier. 2000. Coronavirus: hepatitis, peritonitis, and central nervous system disease, p. 331–348. *In* R. S. Fujinami (ed.), *Effects of microbes on the immune system*. Lippincott Williams and Wilkins, Philadelphia, Pa.
33. Qiu, W., J. W. Park, A. O. Jackson, and H. B. Scholthof. 2001. Retention of a small replicase gene segment in tomato bushy stunt virus defective RNAs inhibits their helper-mediated trans-accumulation. *Virology* **281**:51–60.
34. Rota, P. A., M. S. Oberste, S. S. Monroe, W. A. Nix, R. Campganioli, J. P. Icenogle, S. Peñaranda, B. Bankamp, K. Maher, M. H. Chen, S. Tong, A. Tamin, L. Lowe, M. Frace, J. L. DeRisi, Q. Chen, D. Wang, D. d. Erdman, T. C. T. Peret, C. Burns, T. G. Ksiazek, P. E. Rollin, A. Sanchez, S. Liffick, B. Holloway, J. Limor, K. McCaustland, M. Olsen-Rassmussen, R. Fouchier, S. Gunther, A. D. M. E. Osterhaus, C. Drosten, M. A. Pallansch, L. J. Anderson, and W. J. Bellini. 2003. Characterization of a novel coronavirus associated with severe acute respiratory syndrome. *Science* **300**:1394–1399.
35. Sánchez, C. M., G. Jiménez, M. D. Laviada, I. Correa, C. Suñé, M. J. Bullido, F. Gebauer, C. Smerdou, P. Callebaut, J. M. Escribano, and L. Enjuanes. 1990. Antigenic homology among coronaviruses related to transmissible gastroenteritis virus. *Virology* **174**:410–417.
36. Sawicki, S. G., and D. L. Sawicki. 2005. Coronavirus transcription: a perspective. *Curr. Top. Microbiol. Immunol.* **287**:31–55.
37. Sawicki, S. G., and D. L. Sawicki. 1998. A new model for coronavirus transcription. *Adv. Exp. Med. Biol.* **440**:215–220.
38. Schechter, I., and A. Berger. 1967. On the size of the active site in proteases. I. Papain. *Biochem. Biophys. Res. Commun.* **27**:157–162.
39. Sims, A. C., J. Ostermann, and M. R. Denison. 2000. Mouse hepatitis virus replicase proteins associate with two distinct populations of intracellular membranes. *J. Virol.* **74**:5647–5654.
40. Sola, I., J. L. Moreno, S. Zúñiga, S. Alonso, and L. Enjuanes. 2005. Role of nucleotides immediately flanking the transcription-regulating sequence core in coronavirus subgenomic mRNA synthesis. *J. Virol.* **79**:2506–2516.
41. Teng, H., J. D. Pinon, and S. R. Weiss. 1999. Expression of murine coronavirus recombinant papain-like proteinase: efficient cleavage is dependent on the lengths of both the substrate and the proteinase polypeptides. *J. Virol.* **73**:2658–2666.
42. van der Hoek, L., K. Pyrc, M. F. Jebbink, W. Vermeulen-Oost, R. J. Berkhout, K. C. Wolthers, P. M. Wertheim-van Dillen, J. Kaandorp, J. Spaargaren, and B. Berkhout. 2004. Identification of a new human coronavirus. *Nat. Med.* **10**:368–373.
43. van der Most, R., W. Luytjes, S. Rutjes, and W. J. M. Spaan. 1995. Translation but not the encoded sequence is essential for the efficient propagation of the defective interfering RNAs of the coronavirus mouse hepatitis virus. *J. Virol.* **69**:3744–3751.
44. van der Most, R. G., P. J. Bredenbeek, and W. J. M. Spaan. 1991. A domain at the 3' end of the polymerase gene is essential for encapsidation of coronavirus defective interfering RNAs. *J. Virol.* **65**:3219–3226.
45. Woo, P. C., S. K. Lau, C. M. Chu, K. H. Chan, H. W. Tsoi, Y. Huang, B. H. Wong, R. W. Poon, J. J. Cai, W. K. Luk, L. L. Poon, S. S. Wong, Y. Guan, J. S. Peiris, and K. Y. Yuen. 2005. Characterization and complete genome sequence of a novel coronavirus, coronavirus HKU1, from patients with pneumonia. *J. Virol.* **79**:884–895.
46. Ziebuhr, J. 2005. The coronavirus replicase. *Curr. Top. Microbiol. Immunol.* **287**:57–94.
47. Ziebuhr, J., E. J. Snijder, and A. E. Gorbalenya. 2000. Virus-encoded proteinases and proteolytic processing in the *Nidovirales*. *J. Gen. Virol.* **81**:853–879.
48. Ziebuhr, J., V. Thiel, and A. E. Gorbalenya. 2001. The autocatalytic release of a putative RNA virus transcription factor from its polyprotein precursor involves two paralogous papain-like proteases that cleave the same peptide bond. *J. Biol. Chem.* **276**:33220–33232.
49. Zuker, M. 2003. Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic Acids Res.* **31**:3406–3415.
50. Zúñiga, S., I. Sola, S. Alonso, and L. Enjuanes. 2004. Sequence motifs involved in the regulation of discontinuous coronavirus subgenomic RNA synthesis. *J. Virol.* **78**:980–994.